

Open Research Online

The Open University's repository of research publications and other research outputs

A functional model of similarity

Thesis

How to cite:

Whalley, Peter C. (1984). A functional model of similarity. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 1984 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

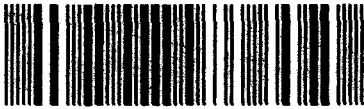
Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000d661>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk



A functional model of similarity

Peter C Whalley

This dissertation is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy to the Department of Psychology of The Open University.

February, 1984

Author's number: HDD 5112

Date of submission: February 1981

Date of award: 19 July 1984

IMAGING SERVICES NORTH

Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
www.bl.uk

BEST COPY AVAILABLE.

VARIABLE PRINT QUALITY

IMAGING SERVICES NORTH

Boston Spa, Wetherby

West Yorkshire, LS23 7BQ

www.bl.uk

**TEXT BOUND CLOSE TO
THE SPINE IN THE
ORIGINAL THESIS**



THE OPEN UNIVERSITY

The Open University,
Walton Hall,
Milton Keynes,
MK7 6AA,

Telephone: Milton Keynes 74066 (Switchboard)
Milton Keynes 6 (Direct Line)

INSTITUTE OF
EDUCATIONAL TECHNOLOGY

Your ref: HDD 5112

Our ref: PCW/sas

20th February, 1985

Mr M.R. Sullivan,
Higher Degrees Office

Dear Mr Sullivan,

I confirm that I am willing for my thesis to be made
available to readers and may be photocopied, subject to the
discretion of the Librarian.

Yours sincerely,

P. C. Whalley,
Research Fellow.

The Open University

Higher Degrees Office

21 FEB 1985

Prints to _____

Disposal: _____

Abstract

In the first chapters, the prevalent conception of the nature of similarity is shown to be too narrow, and the argument is developed that models of similarity must encompass both analytic and synthetic components. Some general problems of measurement and the testing of psychological models are also discussed.

A review of the philosophical treatment of the concept of similarity is made in order to understand the origins of the models found in the present psychological literature. These are then analysed in terms of their implicit object representations and cognitive processes. The distinction between the 'class inclusion' and 'distance relation' models of similarity is shown to be qualitative in nature, and to correspond to the analytic-synthetic distinction in terms of cognitive process.

A functional model of the psychological processes and object representations involved in similarity judgements is then proposed. The fundamental idea involved in this model is that the global properties of referents are synthetically evaluated in terms of their contextual relations, whilst an analytic 'pattern matching' of local properties is made. Various theoretical aspects of the model are examined experimentally, and its general applicability is indicated in a series of applied studies.

The scope of the argument is finally broadened to encompass a development of Torgerson's (1965) conception of the nature of the dimensions resulting from MDS analysis. Dimensions may be considered as 'virtual' artifacts of the experimental task and the individual's conception of it. This possibility allows the methodology to escape the dominating influence of its psychophysical tradition, and become a conceptually deeper tool for cognitive psychology.

To Mary and Richard Whalley

My dear Mary and Richard,
I have been thinking of you both very much lately, and
wondering how you are getting on. I hope you are
both well and happy.

Yours truly,
John Galsworthy

Acknowledgements

I am greatly indebted to my thesis supervisor Jon Slack for his understanding, and patient guidance.

Also I am extremely grateful to Shirley Scripps for her undaunted efforts in the analysis of data and the final preparation of this thesis.

I would like to thank David Neal for helping to devise the pain perception study, and George Wright for providing the facilities that allowed me to carry out the experimental work.

I would also like to thank Richard Austin and the other staff of ACS for their help, and the many members of my own department who have provided support and encouragement.

Finally I wish to acknowledge the contribution of Liz, Ben, Matthew and Jessica in providing the constant reminder that there are at least four things more important to me than finishing a thesis.

Table of Contents

CHAPTER ONE Introduction

1.1 An overview	1
1.2 The parts of this thesis	5
1.3 Measurement	9
1.4 Models	13
1.5 A perspective	17

CHAPTER TWO Similarity

2.1 Overview	18
2.2 Analysis and synthesis	19
2.3 A philosophical framework	20
2.4 Psychological models of similarity	
2.41 Overview	32
2.42 The literature	33
2.43 Distance models of similarity	47
2.44 Content models of similarity	56
2.45 The concept formation literature	73
2.5 Summary	75

CHAPTER THREE A functional model of similarity judgements

3.1 Overview	81
3.2 The conceptual necessity for synthesis	82
3.3 The representation problem	84
3.4 A summary of issues	91
3.5 A functional model of similarity and its consequent process effects	95
3.6 The operation of the model	98
3.7 Experimental tests of the model	101

CHAPTER FOUR Experimental studies

4.1 Overview	103
4.2 Experiment 1: Context effects	
4.21 Rationale	105
4.22 Method	110
4.23 Results	114
4.24 Discussion	124

4.3 Experiment 2: Strategy and complexity effects in similarity judgements	
4.31 Rationale	126
4.32 Method	130
4.33 Results	134
4.34 Discussion	141

CHAPTER FIVE Applied studies

5.1 Overview	143
5.2 The perception of pain	
5.21 Rationale	144
5.22 Method	146
5.23 Results	147
5.24 Discussion	153
5.3 Expert knowledge in a micro-processor course	
5.31 Rationale	154
5.32 Method	156
5.33 Results	157
5.34 Discussion	159
5.4 The perception of argument in text	
5.41 Rationale	160
5.42 Method	163
5.43 Results	165
5.44 Discussion	168
5.5 General discussion	169

CHAPTER SIX Conclusion

6.1 A summary	171
6.2 Prospects and implications	173

References	176
------------	-----

Appendix A	185
Appendix B	188
Appendix C	192
Appendix D	194
Appendix E	197
Appendix F	198
Appendix G	199

Figures

Figure 1.1	Mapping concept relations	7
Figure 1.2	Flow diagram relating inferential classification to the real world	12
Figure 1.3	Flow diagram relating thought to data	12
Figure 2.1	Transformations by different grammars	30
Figure 3.1	A possible network formalism implying local and global aspects	92
Figure 3.2	A comparison of process implications between the 'literary' and 'formal' models of similarity	92
Figure 4.1	Constrained solution under the area (horizontal) and shape (vertical) factorial hypothesis	107
Figure 4.2	Constrained solution under the width (horizontal) and height (vertical) additivity hypothesis	107
Figure 4.3	Schonemann's hypothesis that the perceived shape difference increases as the area increases	107
Figure 4.4	Takane's proposed alternative hypothesis to Schonemann's	108
Figure 4.5	Takane's unconstrained two-dimensional solution for his data	108
Figure 4.6	Dimension weights from an ALSCAL analysis of the 'rectangles' data	116
Figure 4.7	MAXSCAL-4 analysis of the 'skylight' group	121
Figure 4.8	MAXSCAL-4 analysis of the 'mirror' group	122
Figure 4.9	MAXSCAL-4 analysis of the 'rectangle' group	123
Figure 4.10	An example of the cartoon face and bar chart formats	131
Figure 4.11	A 'scenic' strategy being used with the cartoon face format	136
Figure 4.12	A 'groups' strategy being used with the cartoon face format	136
Figure 4.13	Two stages in the use of a 'groups' strategy with the bar charts	137
Figure 4.14	Categorization analysis of ranking strategies: by category, and by subject group	138
Figure 4.15	Categorization analysis of ranking strategies: by dimensionality, and by subject group	139
Figure 4.16	Dimensional weightings for 2, 3 and 4 dimensional stimuli: by subject group	140

Figure 5.1	Two dimensional analysis – burns patients	148
Figure 5.2	Subject weights for the bandage and plastic film burns treatments	149
Figure 5.3	Two dimensional analysis – vertical and horizontal incision patients	150
Figure 5.4	Subject weights after operations involving vertical and horizontal incisions	151
Figure 5.5	MAXSCAL-4.1 analysis of the Incision data	152
Figure 5.6	Two dimensional analysis of the professional programmers' judgements of the micro-processor course concepts	158
Figure 5.7	Dimensional weights of students and tutors	158
Figure 5.8	Dimensional weights for the group reading the pro-discovery method text	166
Figure 5.9	Dimensional weights for the group reading the anti-discovery method	166
Figure 5.10	Dimensional weights for the group reading the anti-discovery method text followed by the pro-discovery method text	167
Figure 5.11	Dimensional weights for the group reading the pro-discovery method text followed by the anti-discovery method text	167

Tables

Table 4.1	ALSCAL sub-group analyses – area-shape hypothesis	117
Table 4.2	ALSCAL analyses of the 'rectangles' data	118
Table 4.3	MAXSCAL-4.1 analyses of the 'rectangles' data	119
Table 5.1	Comparative analyses of the questionnaire data in the burns study	147

1.1 An Overview

"The truth is that Experience is trained by both association and dissociation, and that psychology must be writ both in synthetic and in analytic terms." William James (1890, p487)

The concept of similarity is of fundamental importance to psychology. However it will be shown that an over narrow, and purely analytic, conception of its nature has been adopted in the literature. A general model of similarity, and a better understanding of the nature of the cognitive processes involved in similarity comparison tasks, is developed. The essential argument of this thesis is that similarity judgements must be seen as incorporating aspects of both analysis and synthesis. It is also found to be necessary to examine the implicit representation systems of the existing theoretical models of the psychological literature.

The philosophical literature has always paid some regard to both the holistic and analytic aspects of similarity, and in some languages there may still be found clearly distinguished expressions. For example the two Sanskrit terms for similarity, *sadrsya* and *sadharmya*, explicitly refer to 'way of being' and 'common property' relations respectively. The early psychological literature also allowed for other than a purely analytic perspective. Herbert Spencer defined a form of 'imperfect similarity' in terms of the "likeness of nature in the relations" of objects (1881,p259). William James also attempted to examine the representational implications of the judgement process,

"Is the 'resemblance' which we predicate of two objects due in the last resort always to the operations on our mind of qualitatively identical elements contained in each? Or, may we, on the other hand,

admit the existence, amongst our minds's objects, of qualities or natures which have no definite 'point' in common, but which we perceive to be, although numerically distinct, yet like each other in various degrees and ways?" (1893, p333)

The various models of similarity judgements in the present psychological literature may be usefully dichotomized as being concerned either with 'content' or 'dimensions'. This distinction was first proposed formally by Ekman and Sjöberg (1965), and will later be shown to correspond to the operation of analytic and synthetic cognitive processes. The conceptual distinction between the forms of models is illustrated in Attneave's widely referenced review paper:

"It is obvious that when things are similar they are similar with respect to something. The characteristics with respect to which objects are similar may be conceptualised either as more or less discrete and common elements or as dimensions on which the objects have some degree of proximity." (1950, p519)

In the dimensional or 'geometric' models, similarity is considered to be the complement of distance in a space of one kind or another. The individual's cognitive structure is implicitly conceived of as objects distributed within a 'psychological space'. The origins of these models may be traced back to the early Gestalt psychologists, and conceptually they have not developed from the original concern for psycho-physical phenomena. The reason for their dominance of the psychological literature was the early development of the multi-dimensional scaling (MDS) algorithms (Shepard, 1962). This meant that a wide range of problems could be tackled with easily realised data. However in a recent, and extremely influential paper, Tversky (1977) has challenged the dimensional-metric assumptions that underlie the geometric approaches to similarity, and developed an alternative 'feature-theoretic' or 'contrast' model. His model may be characterised as being of the 'content' type, and a detailed review of both the claims made

for this approach, and their refutation by other authors, will be made later. The tenor of his argument may be gained from the following:

"- the similarity between two objects is expressed as a feature-matching function (i.e., a function that measures the degree to which two sets of features match each other) rather than as the metric distance between points in a coordinate space. - the similarity of a to b is described as a linear combination (or a contrast) of the measures of their common and distinctive features. Naturally, similarity increases with the measure of the common features and decreases with the measure of the distinctive features."

Tversky (1977, p79)

A qualitative distinction between the content models with their concern for 'class inclusion', and the 'distance relation' models can be made in terms of their implicit analytic and synthetic cognitive processes. Synthesis can be regarded as the search for resemblance, whilst analysis is concerned with the isolation and disjunction between elements of the objects. The appreciation of both of these aspects of similarity estimation has a long history, although in both everyday speech and the present psychological models, the analytic tradition is dominant. We shall argue that this distinction is central to an understanding of the reasoning behind the different models, in that it has implications for the implicit representation of the individual's personal meaning for concepts. Underlying the various models of relational judgements are quite different conceptions of the representation of knowledge, imposing distinct constraints on the possible cognitive process involved. Furthermore it will be shown that the experimental material, choice of knowledge domain and task constraints may also have important effects. It is found that these may cause subjects to switch to more or less appropriate styles of thought, and significantly change the outcome of judgements that they may make.

The essential weakness of the early psychological models which carried some implication of holistic or synthetic process, is that they make an almost immediate reference to atheoretic innate or metaphysical faculties. For example, Binet wrote critically of the reductionist approach to similarity, but in order to avoid the epistemological traps noted he merely writes, "- we repeat that similarity is a simple final and irreducible notion" (1907, p117; translation from Reeves, 1969, p189). This thesis is an attempt to develop the ideas of Spencer and James within a framework of sufficient correctness to bear comparison with the formal models of the present day. The development proposed here is a greater understanding of the **necessity** to postulate both analytic and synthetic processes in similarity judgements. This is made in terms of the 'qualitives' and 'relatives' of the objects being considered; judgements of the 'qualitives' being related to analytic, and the 'relatives' to synthetic thought processes respectively. These terms, derived from the philosophical literature are related to the equivalent expressions of the cognitive psychology literature, i.e. Gardner's (1978) 'content' and 'configural' aspects of the stimulus. A more general model of comparative judgements is put forward, and proof of the central argument, and the various developments, is offered in terms of a series of experimental and applied studies. In these experiments manipulations are made in terms of complexity and general context, and various measures are consequently developed of the resulting cognitive processes.

Part of our general argument will be to indicate the problems associated with attempting to define objects merely in terms of discursive lists of attributes; an important consequence being the inability to deal with their configural and relational aspects. Viewing, and attempting to make a definition of the similarity relation in terms of a process model, appears to be the most satisfactory way to deal with the possibility of such emergent properties. Following Torgerson (1965), the argument is also put forward that experimentally elicited judgements of similarity should be viewed not as a basic, intuitive or perceptual, relation but as a

derived cognitive relation. Torgerson claims that whether the essential representation of objects is multi-attribute or multidimensional, it is always possible to conceive of the structural aspects of the problem space as being embedded in an appropriate space. A consequent development that we shall later consider is that the 'dimensions' resulting from an MDS analysis, are most usefully conceived of as being no more than derived functions of the implicit representations and manipulations of the experiment. In this role they may be seen as measures of emergent configural effects. These cannot be accommodated within the simple feature-theoretic framework, and the forms of dimensional analysis appear best suited to their systematic evaluation.

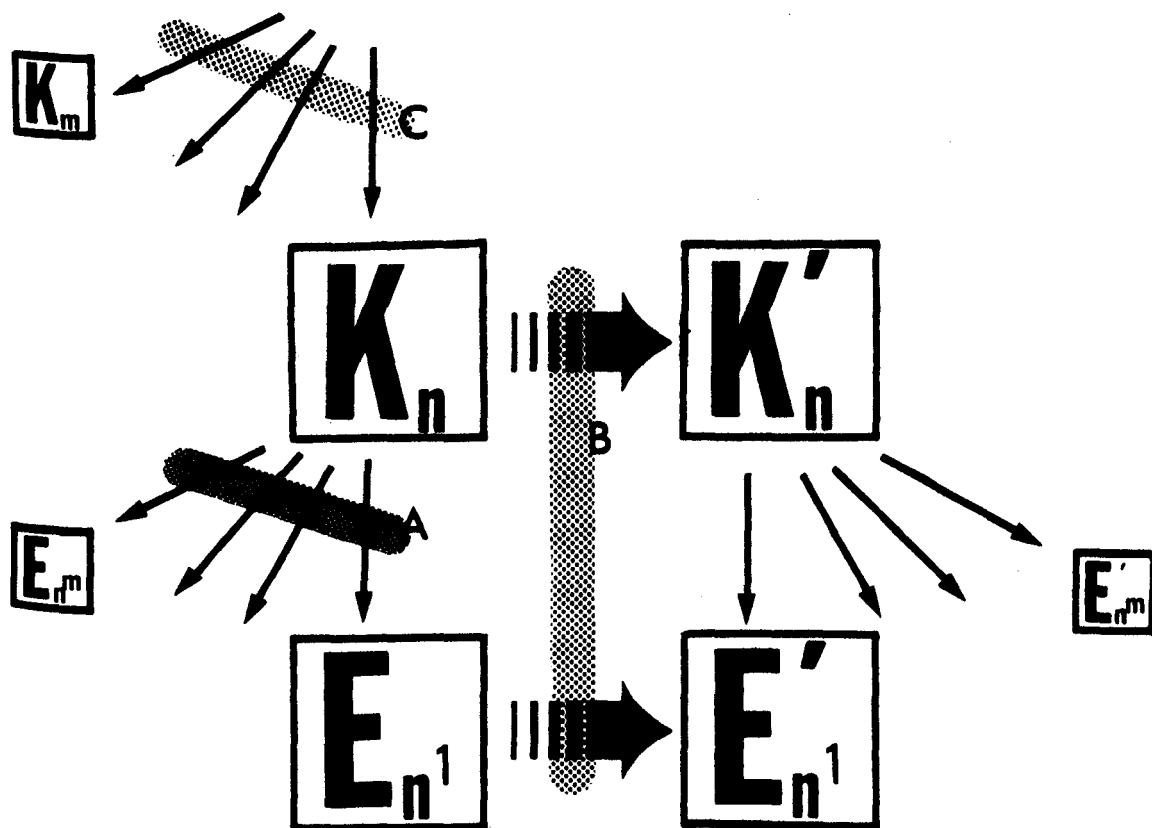
1.2 The parts of this thesis

The original motivating force behind this thesis was the wish to validate a theoretically and empirically sound methodology with which it might be possible to make investigations of the growth of human knowledge. Such studies were to be in subject domains of 'knowledge with belief'; where the values, and perhaps prejudices, of the individual must be taken into account and one could not expect consensus of view. Furthermore the development of knowledge was to be viewed as more than the mere 'accumulation of more content'. The basis of any psychological investigation of the growth of knowledge is the methodology adopted for measuring change. For the topics chosen, it was found necessary to map the individual's concept relations at intervals in the learning process. The analytic procedures available to the researcher interested in mapping concept relations had generally evolved from the concerns of the psychometric community to obtain 'hard', in the sense of being statistically sound, data. An unfortunate consequence of this quite legitimate desire, were methodologies involving complex relational judgement tasks that may take place over many hours and often involve

completely context free judgements; these constraints being imposed in an effort to reduce inconsistent or intransitive responses. Psychometric texts invariably deal with the 'meaningfulness issue'. However, this is always taken to refer to the scale or dimension which is being created not the meaningfulness of the task to the subject; the source of the data upon which so much analytic effort is later expended.

An essential component of procedures for making mappings of the internal representation of knowledge is some form of relational judgement between items in a knowledge domain. The emphasis of this thesis therefore became directed towards the theoretical issues underlying similarity estimation. A set of small scale applied studies are developed later to provide some indication of the soundness and generality of the model of comparative judgements proposed. It is intended that the presence of these applied studies should not confuse the development of the main argument, although it will be apparent that their general nature has strongly influenced the direction and development of the theoretical parts.

Figure 1.1 represents an attempt to make clear the relationship between the individual's world knowledge and the possible 'mappings' of portions of this knowledge. Our intention here is to delimit the main theoretical issues of our argument from other issues that are implicitly related, but only of importance in an applied context. Changes in the inferred knowledge domain K to K' might involve providing an individual with more or different information on a particular topic. It could also take the form of a group difference in perception or experience. The mapping from K to 'elicited' knowledge E is the product of some testing procedure, and focus A therefore deals with the validity of the mappings provided by the various elicitation procedures. The tasks studied within focus A involve the learner in making relational judgements and are therefore implicitly dependent on some model of similarity. It is this dependency and the theoretical consequences of the validity and accuracy of the elicited mappings of the individual's knowledge that is the main concern of this thesis.



Key
 knowledge domain
 elicited knowledge
 relational system
 'mapping' relation

K
E
 □
 →

change in relational
 system over time

main locus of study

subsidiary locus of study

→
 —
 ···

Figure 1.1 Mapping concept relations

The issues involved in focus A are dealt with in a set of theoretically orientated experiments in Chapter 4. These demonstrate the necessity for a more general model of similarity estimation. Although a particular underlying model can be shown to be theoretically more general or adequate in some way, it is necessary to look to applied studies to give some context to these differences which are necessarily purely qualitative. Thus changes in focus B might be viewed as a measure of 'adequacy of explanation', and the 'power to generalise' of focus C as an external measure of validity. The set of applied experimental studies in Chapter 5 are intended to cover some of these issues. For example if it can be inferred that an individual has a known level of precision of definition at level K then the elicited level E is inadequate in some way if it is not also capable of such differentiation. It is also preferable that an elicitation technique should meet this criterion of adequacy in as wide a range of topic domains as possible, and not be restricted to subdomains of experience; from which generalizations of only doubtful validity may be made.

The effects on the individual of change over time, restructuring and learning would be contained in focus B. A text manipulation experiment concerned with the role of play and mathematics in the primary school was included to test for the envisaged effects of 'knowledge with belief'. Focus C is concerned with the problems of 'creating' sub-domains of knowledge, and with the ability to generalise of the various testing techniques. An attempt was made to tap the 'growth' of knowledge of individuals; and to cover a range of knowledge domains. A study of the concept relations involved in a micro-computer course was included to cover aspects of descriptive knowledge, and a domain where language could be presumed to play only a minimal role, the perception of painful experiences, to represent aspects of knowledge best considered in some non-declarative form.

1.3 Measurement

"Traditionally psychometrics has had its fair share of simple nonsense; now like any other branch of applied mathematics, it is in danger of having its share of complicated nonsense." Gregson (1975, p186)

Measurement as a problem.

Gregson is critical of the contentment within the psychometric tradition with attempts to fit mathematical equations to psychophysical functions, rather than seeking to understand the processes involved. Historically this desire to remain outside the 'black box' may be seen as a carry over from the rodent experiments and the subconscious physiological processes tackled in the 'building bricks' approach to psychology. For example Lingo takes the contrary view that Fechner's mathematisation of Weber's law, "laid the foundation for all the important methods of psychological measurement", (1979, p5) in that it represented the first successful attempt to describe a psychological process in functional terms. Lingo regards measurement as being the most "pervasive and fundamental problem for any discipline that aspires to be a science" (p1). Furthermore he takes the view that the "sophistication, generality and power of a science can be judged on the basis of its measurement procedures and its use of mathematics" (p2). The general claim is made in this thesis that the detailed concerns of experimental psychology have often led to a quest for mathematical accuracy at the expense of psychological validity. It could be argued that the present 'zeitgeist' of context dependent psychological studies might well be only a temporary phenomena, but there is no a priori reason why they should be invalidated by the lack of suitable analytic techniques.

The nature of error

As discussed previously our main emphasis will be directed towards an evaluation of the validity and accuracy of similarity judgements. The detailed analysis of such relational data can become fairly involved, but

essentially takes the form of a dimensional or hierarchical analysis. The actual form of the analysis is bound up with the implicit model of judgement, but this too is to be developed. A comparison of such analyses is made in Chapter 2, and the various techniques developed for determining which algorithm best 'fits' the data are outlined. However it will be noted that these algorithms optimise their final configurations on quite **different** aspects of the data. Thus ruling out strong claims as to the possible processing strategy or knowledge representation, of individuals in the original experimental setting. The dimensional scaling analyses make most use of information concerning those objects which are **least** similar whilst the hierarchical cluster analyses are dependent on information about the **most** similar objects. Claims as to causality would be significantly weakened if the individual is found not to produce errors in the form of Gaussian noise, but is being consistently biased by some artifact of the experimental design or setting.

Another important form of error, **intransitive** judgements, is conventionally treated in a similarly atheoretical fashion. Despite the possible psychological significance of such relations, they are either discarded or 'averaged out' by all the generally available algorithms; where the experimental procedure does not already constrain their occurrence. Explicit predictions concerning the nature, source and implications of 'error' should therefore be included in any theory of relational judgements. These problems are dealt with after our own model of relational judgements is developed. One hopeful sign in this context is the 'maximum-likelihood' algorithms as developed by Takane (1981), and used in the later sections. These permit the investigator to make and test theoretical predictions concerning the way in which similarities data is **generated** in different contexts. This will obviously be a powerful tool for the concerned investigator, particularly when these algorithms are extended to all forms of relational judgement data.

Making inferences from the real world

"Quite generally (despite the cant of some of the more hardnosed philosophers of science), we can, within certain limits, paraphrase Humpty Dumpty to mean by measurement whatever we choose it to mean, sometimes more and sometimes less, depending upon what purpose we have in mind." Lingoes (1979, p6)

A synthesis, or at least the source of disagreement, of the contrary views concerning measurement can be found in the way that Coombs distinguishes between the two common uses of the term 'data' in the social sciences. It may refer to either the recorded observations that the scientist has chosen to observe or to "that which is analysed" (1964, p4). Coomb's representation of the 'scientists' inferential process is shown in Figure 1.2, and it is interesting to note the different emphasis accorded to each phase. From the universe of potential observations the scientist **selects** 'some few things to record in Phase 1. An **interpretive** step in Phase 2 involves a 'classification of observations'. However the optional decision as to which 'kind of data' the observations represent is adjudged to be the **creative** step in 'collecting the data'.

Coombs discusses the restricting effects on lower phases in his model of decisions made at a higher level, but, does not mention the limiting effect on the higher levels that a dominant model of 'data' and the availability of analytic techniques might induce. Coombs' work of course is primarily concerned with the different kinds of data and their interrelations, and Figure 1.2 would appear to represent a model emphasising the collection of data **acquired for its own sake**. A 'problem driven' model of the data acquisition process is indicated in Figure 1.3. The goal of the experimenter is here taken to be in some way concerned with elucidating the individual's thought processes. To achieve this, a subgoal of a judgement task is introduced which may be considered as a **constrained** mapping of K. Also a data elicitation procedure is set up, with the data generated taken to be a

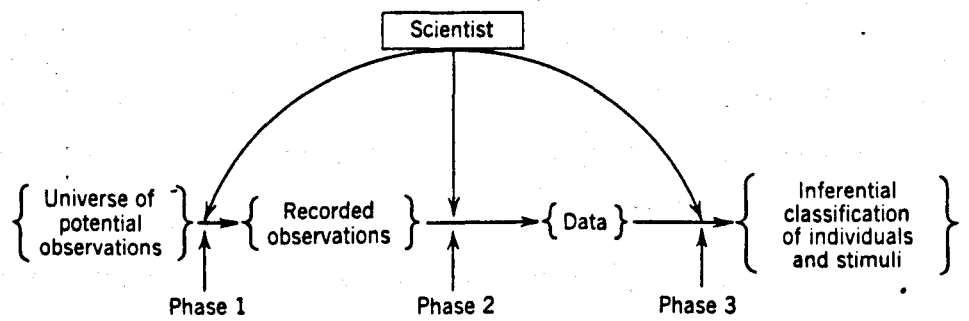


Figure 1.2 Flow diagram relating inferential classification to the real world.
(from Coombs, 1964)

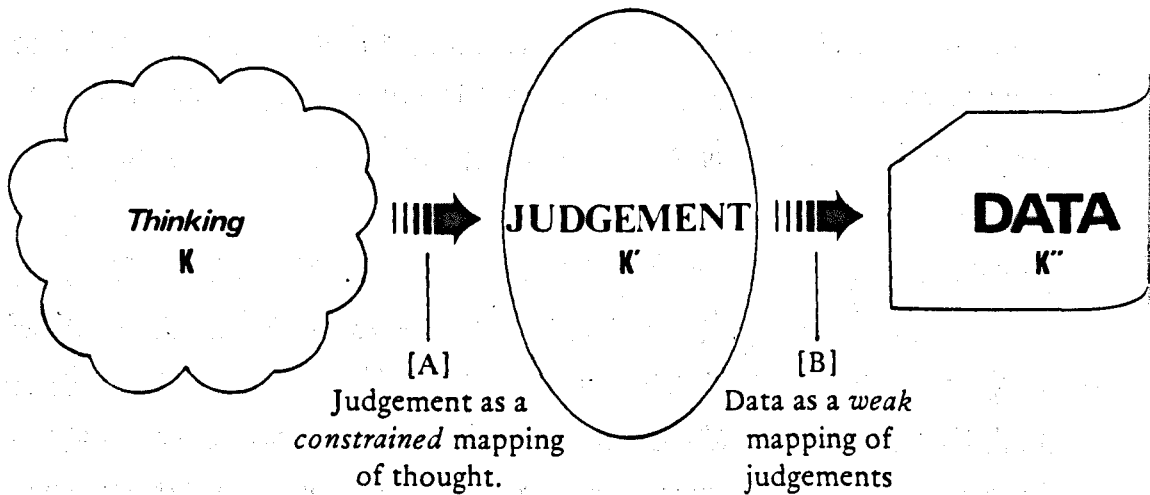


Figure 1.3 Flow diagram relating thought to data.

weak mapping of K'. The contention is that the conventional psychometric concern has been with the accuracy of mapping [B], that this has led to artificially constrained judgement tasks, and to mapping [A] being effectively ignored. The intention of this thesis is to move the level of concern back to the primary mapping and to evaluate the consequences of accepting psychometrically 'weaker' data in a less constrained, and therefore psychologically richer, situation. The implications of making contextual judgements of complex entities are discussed in Chapter 3 in terms of a general model of relational judgements, and in Chapter 4 in terms of the necessary changes in data collection and analytic techniques.

1.4 Models

The required accuracy of models

The social sciences often appear to have an uncertain perception of the worth of the conceptual framework employed by the natural sciences. At times it is dismissed as inappropriately naive 'scientific absolutism' and yet at others viewed as the touchstone of 'proper' methodology. Putnam develops the idea that models associated with the 'higher-level laws' of psychology need only to be true within the accuracy of psychology and that "the model associated with higher level laws need not **at all** be compatible with the model associated with the lower-level laws." (1981,p214). He concludes this argument with the suggestion that "**oversimplified** models may well be the best possible at the **higher levels**" (p214). The implication of this reasoning is that experimental exploration in the higher level sciences such as psychology must resemble a satisficing process, optimising descriptive detail with explanatory power at a psychologically interesting level. It is possible to contemplate a 'utility theory of measurement', with a U-curve relationship between psychological interest of the outcome and the detailed accuracy of the data. For example it will be claimed in Chapter 3 that an

apparently less parsimonious 'process' model provides a realistically more detailed insight into the way in which similarity judgements are made, than the equivalently more detailed but 'static' feature analytic models.

Pylyshyn argues just this point in advocating the pursuit of 'top-down' research strategies rather than the traditional 'building-bricks' approach. He contrasts the empirical and rational sources of understanding with their "two corresponding loci of rigor". (1979, p427) If the researcher's intuitions and personal knowledge are not employed in rigorously ensuring "that one's theoretical ideas are complete, consistent, and logically sound" (p427), then the consequence of rigorous operational definition of constructs and experimental control may still be incoherent and contradictory theories. It is claimed that the researcher must achieve a balance between the precision of isolated theories ("mechanisms") which might account for a limited set of experimental results and the "deeper, more general and therefore more **complete**, explanation" (p428) to be gained from a 'top-down' approach, such as the functional analysis implicit in viewing complex systems as interconnecting 'black-boxes'. It is unfortunately the case that most of the theories of similarity discussed in later chapters consist of little more than operationally defined properties and parameters. They are able (and were so intended) to 'predict' experimental results, "while postponing the explanation of how these properties come about, or what more general role they might have in cognition". (p428)

Reductionist arguments

"We often think that when we have completed our study of one we know all about two, because 'two' is 'one and one'. We forget that we still have to make the study of 'and' - that is to say, of organisation."

L von Bertalanffy (1973)

It will be argued in the next chapter that the primary distinction between the various models of similarity judgement is really at the 'theory of science' level, and stems from the reductionism implicit in the feature analytic models, such as Tversky's. Such models have been criticised for being too prescriptively normative. However as with the formal AI models in psychology they appeared to be 'powerfully' explicit, and it has been taken that if they fitted the scaling data even only slightly better than the more general dimensional models, that they were therefore more correct.

A general problem for any reductionist argument is that the proponent is obliged to provide an a-priori explanation of the decomposition imposed, and to prove that the inherent simplification involved in the reduction will still leave an adequately descriptive model of the original complex. Also it is necessary to show that the conventionally simplistic additive rules employed provide a valid recomposition, whose product is formally equivalent to the individual's decision processes. The contention here is that the feature models have to prove that they have not just artifactually optimised a few features of the experimental setting, and that they therefore could **always** be tuned to give 'better' results. Katz and Fodor's (1963) ideas concerning a feature analytic theory of word meaning similarly appeared to work well when dealing with simple concrete concepts, but are extremely difficult to generalise to more abstract ideas; whose meaning is derived from a pattern of relations built up from their use in context.

The grain problem

The other fundamental flaw to which reductionist models are liable is that they are likely to represent, and be concerned, with the wrong level of organisation of the data. There is a means-end reversal common to much cognitive psychology theorising, of delimiting the world to manageable proportions and then expecting to be able to generalise to more complex situations. The more reductionist the model employed, the more likely the theorist is to lose track of the **purpose** of things and merely indulge himself in precise but meaningless quantification. The effects of such 'data driven' research frameworks was mentioned earlier, and it is a common theme within this thesis that the dominance of certain data elicitation techniques has effectively circumscribed the form of later studies.

Putnam argues that current 'data flow' representations, such as production systems within the Artificial Intelligence paradigm, are good examples of this problem. If AI is seen as concerned with the organisation of cognition, then production systems, even when they can be made to appear to work, are psychologically uninteresting in that they are a formal description of the wrong level of description of the data. Indeed it has been suggested that the hidden command structures that are required to make such systems 'work', are their sole psychological point of interest.

Breaking models down.

The inability to distinguish between essential and accidental features of a model may result from the problems outlined above. However as Gregson points out in the context of theories of similarity judgements, the most important consequences for psychological theorising arise when such models **break down**, "The most complex set theoretic notions (of similarity) provide no more information than the operationalists notions of response identity." (1975, p10). The essence of a good analogy (or psychological theory) is the amount of information it provides when disconfirmed. Essentially normative models that are little more than data fitting

exercises, and in addition severely delimit the field of experimentation, are not the obvious way to acquire new knowledge.

1.5 A perspective

This chapter has been intended to highlight the implications for theories of relational judgement of the general move towards more realistic, perspective-dependent experimentation. A particular concern of this thesis is the use of such judgements in attempts to 'tap' the individual's knowledge states. If it is accepted that the individual will attempt to make use of prior knowledge and will be influenced by the general context of experiments, then any complete model of human judgement must make explicit predictions concerning the effects of such 'top-down processing'. The contention is that for the more complex domains which are tackled in the chapter detailing the applied studies, it is important to make a proper evaluation of the psychometric techniques employed to gain insight into such states of knowledge.

A detailed treatment of the concept of similarity is made in the next chapter. A philosophical perspective is employed in an attempt to clarify the confusion in the psychological literature, and an alternative to the presently dominant feature matching models is developed. This model is intended to more adequately deal with the complex, 'synthetic' relational judgements which are the likely consequence of realistic experimentation. Various theoretical aspects of the proposed model are examined in Chapter 4, and its general applicability is then indicated in a series of applied studies in Chapter 5.

2.1 Overview

"Assuming that we can - get rid of all universals except similarity, it remains to be considered whether similarity itself could be explained away."
Russell (1940,346)

Philosophers have always regarded similarity as a fundamentally important relation. However they have tended to make use of the relation in terms of theories concerning universals and the nature of reality, and not to have addressed it as a problem for investigation in its own right. As might be expected, psychological theorists have shown more concern for how the individual actually makes such judgements, but have generally adopted a rather narrow conception of the nature of the relation itself. Various terms introduced in this chapter are to be developed further in a latter attempt to address the representational issues that underlie the perception of similarity. However before attempting to set up a framework for a review of the psychological literature, it is necessary to provide 'working definitions' of the key terms **analysis** and **synthesis**. The philosophical framework constructed here does not pretend to scholarly depth or detachment, in that it is primarily directed towards an understanding of the origins of the different models of similarity found in the present psychological literature. In the later review of the psychological literature, we shall examine the important distinction made between the 'class-inclusion' and 'distance-relation' models of similarity. The primary purpose of this review is to indicate the source of this distinction, to show that it is truly qualitative in nature, and that it corresponds to the analytic-synthetic distinction in terms of cognitive process.

2.2 Analysis and synthesis

The term 'analytic' is used in the psychological literature to refer to forms of cognitive processing involving resolution into simple elements. For instance, Brooks (1978) writes, "Let us take **analytic** concept identification to mean a process whose direct effect is to separate aspects of the stimulus and evaluate their ability to predict category membership" (p180). The contrary term 'synthetic' is defined in terms of building up separate elements into a connected whole. Brooks consequently regards non-analytic concept formation as resulting from some inferred 'overall similarity'. It should be noted that within the literature, 'synthetic', 'non-analytic' and 'holistic' are usually regarded as equivalent. Psychologists have generally adopted this natural language distinction between analytic and synthetic cognition, whose actual origins lie in the concern of the early philosophers with methods of argumentation. Aristotle's 'Prior Analytics' is concerned with the development of a theory of **inference** and the 'Posterior Analytics' with a theory of **proof**. He contended that we have a **demonstrative** reasoning capacity that might act as a counterweight to our **dialectical** pattern of reasoning. There is a clear correspondence between this dialectic-demonstrative distinction in reasoning and the analytic-synthetic distinction in judgement process.

Within modern logic the terms have become more narrowly defined, and analysis is now identified with 'induction' and synthesis with 'deduction'. Following Kant, philosophers conventionally use the terms 'analytic' and 'synthetic' to refer to **judgements** concerning the truth or falsity of **statements** being a logical or a factual matter respectively. It is probably known to the reader that there are various problems underlying their use in this way. However for the moment we shall only attempt to consider them in terms of a possible processing distinction within judgement tasks. It is intended to discuss the conceptual problems in Chapter 3, when attempting to address the various representational issues that underly the main argument. Within this simplifying context then, analytic judgements

can be defined as making clear the meaning of their terms, but not going beyond them. On the other hand synthetic judgements are about something which is, has been or will be the case; they are practical contextually bound judgements.

2.3 A philosophical framework

Introduction

We have already noted that philosophers have treated the concept of similarity as an important, but conceptually difficult topic. The usefulness of their treatment is sometimes not immediately apparent from a psychological perspective. For example, philosophers have even appealed to the indefinability of similarity to prove the existence of other complex ideas. Butler (1910) claims that although we cannot define similarity (or equality) we know what they are by 'acquaintance with instances'. The existence of the similarity relation is then used to prove our awareness of the existence of personal identity. This is experienced "Upon comparing the consciousness of oneself or one's own existence in any two moments -". As it is difficult to exactly delineate the impact of one philosopher's ideas upon another, the various works are detailed in historical sequence and only the correspondence to our own argument is noted.

Aristotle

Guthrie (1967) describes the Greek philosophers as struggling to express new and difficult conceptions. He points out the problems that "- Aristotle's logic only overcome with difficulty because it is rooted deep in the Greek consciousness: even their language has only one word for **same** and **similar**" (p230). This is evidenced in Aristotle's **Metaphysics** where

although 'other' is treated as being the opposite of 'same', he writes

"The same has several meanings; (1) - the same numerically - (2) - one both in definition and number - (3) - the definition of its primary essence is one -" (1054b).

This last case, the **matching of qualities**, is effectively what Aristotle means by similarity and is later expanded,

"- things are called like if the qualities they have in common are more numerous than those in which they differ: either the qualities in general or the prominent qualities; e.g. tin is like silver, qua white, and gold is like fire, qua yellow and red" (1054b).

It is interesting to note that other translations of this same work use the term 'similar' in place of 'like'. Aristotle's other discussion of similarity in this sense, that is now thought of in terms of 'class inclusion', is to be found in **The Categories**. Here he attempts to differentiate amongst forms of substance, employing the terms 'quantity', 'quality' and 'relatives':

"- most distinctive of a quantity would be its being called both equal and unequal." (6a26)

"Relatives seem also to admit of a more and a less. For a thing is called more similar and less similar, and more unequal and less unequal; and each of these is relative -" (6b19)

"- it is virtue of qualities only that things are called similar and dissimilar." (11a15)

Aristotle's treatment of concepts such as qualities and relations is embedded within Western language and thought. In our later development of these terms it will become apparent that Aristotle's discussion is more lucid than anything to be found in the modern psychological literature.

The alternative to 'class inclusion' is the 'distance relation' model of similarity. The ideas behind this model will be developed later in this section and also in Section 2.4, and our main argument is that the distinction between these two models corresponds to the analytic-synthetic distinction in terms of cognitive process. In Chapter 6 we shall attempt to show how the distance-relation models, and synthetic cognitive processing are related to the use of metaphor in natural language. Aristotle's work is important in the context of this argument both because his views on metaphor have been so influential, and also because of the nature of his perceptions of the process.

"- the greatest thing by far is to be a master of metaphor. It is one thing that cannot be learnt from others; and it is also a sign of genius, since a good metaphor implies an intuitive perception of the similarity in dissimilars." (The Poetics, 1459a)

The Sanskrit tradition

The Indian philosophers of the equivalent era would appear to have had a quite different problem. To them, the perception of differences did not seem to have any intrinsic validity, because of the inherent difficulties which follow from the notion of negation that it implies. These philosophers have tended to take similarity as a primitive relation generated by our conceptualizing activity, which projects our classifications into the world. As with most schools of philosophy, the notion of similarity is found to be intimately bound up with the notion of 'universals', in the sense of abstract generalities:

"There is not, however, much difference between Ramanujacarya's definition of universals and Venkata's definition of it, for though the former defines it as any assemblages that are similar and the latter as similarity, yet the very conception of similarity of Venkata involves within it the assemblage of parts as its constituent; -" (Dasgupta, 1968, vol III p355).

Deshpande (1972) details the fundamental importance of the concept of similarity in Indian poetics, and shows how it forms the basis of many classifications and definitions. The two most important terms used in the Sanskrit works to denote the relation of similarity are **sadharmya** and **sadrśya**. These are conventionally abbreviated to SDH and SDR respectively. A Sanskrit philosopher of the 12th century, Mammata Acharya defined SDH in formal terms as a 'relation of the object of comparison and the standard of comparison to the common property'. In contrast, he regarded SDR as a 'relation of the object of comparison with the standard of comparison', and considered it to be caused by or based on the common property. Deshpande rewords this definition in 'plain language' as,

"it may be said that the fact that two entities have some common property corresponds to the relation of SDH, whereas the cognitive result that those two entities appear similar to each other on account of that common property corresponds to the relation of SDR". (p23)

In some sense the relation of SDR is based on SDH, but is considered to be different in nature. The relation of SDR involves the cognitive factor of samana-darsana 'same or similar appearance', and can be distinguished from SDH in that for SDR the observation of common properties is only a necessary preliminary stage in having a synthetic vision of similarity:

"In this vision, which is a state of cognition, the two entities appear similar as two wholes, irrespective of the differences they have" (p24).

Entities may have some common properties, and also some characteristically individual 'non-common' properties. The relation of SDH corresponds to the common properties, while the opposite relation of VDH (vaidharmya, the 'relation of having non-common properties') corresponds to the non-common properties. In that they do not depend on any subjective factors, SDH and VDH can logically be considered as the direct complement of each other. However the realisation of common properties is not the only causal factor in the case of SDR. The term corresponding to SDR, VDR (vaisadrsya, the 'relation of having dissimilar appearance'), is therefore **not** its direct complement. This is a consequence of the fact that although the cognitive relation of SDR is a product of the realisation of common properties, it does not only consist of such common properties. Deshpande describes SDR as being synthetic in that

"it is of a generalised nature in which the data of common and non-common properties are de-emphasized, and are psychologically synthesized to give rise to a uniform opinion or judgement" (p24).

He makes further use of the conceptual distinction between SDH and SDR to examine the various other terms that may also signify the similarity relation, and also to look at other relevant judgement processes, such as making comparisons. However for our purposes it is most appropriate to conclude with a distinction that he draws in terms of analysis and synthesis, the key terms of our exposition:

"Thus SDH is analytical in character, while SDR is synthetic. While SDH and VDH are derived through collection of data, SDR and VDR depend upon the choice and taste of a person. SDH is a logical relation, while SDR is a cognitive or psychological relation." (p26)

Boethius of Dacia

Skipping to the 14th Century, and a continent, we find Boethius noting the effects of **context** upon similarity. From the time of the Greeks,

philosophers concerned with the validity of rhetoric have criticised the use of 'argument by analogy'. However Boethius explicitly detailed this as a problem in terms of making similarity judgements. Within the framework of an analytical model of similarity, he showed that the common concepts on which the 'class match' were being made, are a function of relative context; and used this argument to explain various forms of rhetorical error. Boethius was concerned with the possible definitions of entities, and their use in arguments. He considered things to have 'certain modes of being'; from which common concepts might be drawn. Such common concepts are relatives and

"- because a common concept is only **accidentally related** to the thing of which it is a concept, the Topic drawn from it is not a cause of the conclusion of a dialectical syllogism".

Peirce

The 'softness' and changeability of language are conventionally accepted as a condition for its development. Analytical philosophy translates the question of vagueness into the extensional-intensional problem of meaning. However, to Peirce, vagueness is a question of representation and not some peculiarity of the object being represented. Vagueness is defined as an implicit part of any sign process, and thought to constitute a universal principal. Peirce would therefore claim that vagueness is not the result of a defect in thinking or knowledge, and developed the notions of 'breadth' and 'depth' to distinguish it from ambiguity; the logical deficiency of meaning. Underlying Peirce's philosophy is the 'doctrine of continuity', or synechism, which corresponds to the Sanskrit philosophers' concern with 'unity'. It will be necessary to return to Peirce's ideas when we attempt to address the representation problem. However the importance of his work for our present argument is that, as Nadin (1980) points out, his ideas concerning the continuity and process nature of reality were later taken up by Wittgenstein in his idea of the **gradual nature** of similarities.

Russell

It will be noted later that Russell's work is the key philosophical reference for the class-inclusion, or in our terms analytic, models of similarity in the psychological literature. Whilst attempting to set up a definition of cardinal numbers he writes,

"When two classes have the same number, they are said to be similar"
(1903, p113)

and later,

"We have already seen that, if two classes be called **similar** when there is a one-one relation which couples every term of either with one and only one term of the other, then similarity is symmetrical and transitive, and is reflexive for all classes" (p305).

This definition of similarity in terms of class inclusion is that adopted by those psychologists who have chosen to use set theoretic terms to describe their theories, or have fitted their theories into the formalism of set theory. It is clearly not original to Russell, although he is the source cited in the psychological literature. What is missed in these references is Russell's treatment of the **problems** that follow from this definition of similarity as a form of partial identity; namely those of **context** and **relevance**. He showed that if similarity is understood to be a matter of degree, then the statement that things similar to A are always followed by things similar to B may not be true, even in cases where A is the cause of B. This 'causality problem' is illustrated by the case of two pairs of matches. The pairs are taken to be 'similar' to each other in all respects except for colour in the first case, and dampness in the second. The degree of similarity between the members of each pair is identical in that they differ in only one respect, but quite different in respect to whether the matches will ignite in the same way; a matter of the relevance of an aspect within a particular context.

Russell's work is important in respect of the development of both classes of models of similarity to be found in the psychological literature. In the same volume that the concept of similarity is defined in terms of, or perhaps more accurately is used to define class relations, Russell also develops the notion of similarity as a **distance relation**. In attempting to distinguish between qualitative and quantitative comparisons, he describes 'sweetness' and 'brightness' as genuinely quantitative but regards terms denoting colour such as 'ruddiness' as indicating,

"- not more of a given colour, but more likeness to a standard colour. Various shades of colour are supposed to be arranged in a series, such that the difference of quality is greater or less according as the distance in the series is greater or less" (p171).

Russell describes the idea of an 'ideal ruddiness' and notes that for all such colour terms, "The true quality involved seems to be, in all these cases, a relation, namely the relation of similarity" (p171). Russell regards the difference between shades as one of **quality** rather than just of magnitude. However in that the difference of quality is itself a matter of degree, this is a case of relations which have magnitude; and the relation between any two of them may "in a generalised sense be called a **distance**" (p171). Russell notes that each shade of colour appears to be simple and **unanalysable**, and describes such relations as being of 'immediate resemblance'.

Again these ideas concerning distance relations are not original to Russell, he references the work of the early German psychologists that we shall review later. However his analysis is important in that he explicitly relates the concept of similarity to distance relations, and also because his ideas appear to have influenced Torgerson; one of the founding fathers of the 'scaling', or distance relation, model of similarity judgements. We therefore have the apparent paradox of Russell being referenced as an original source by both opposing camps!

Wittgenstein

We have already noted Nadin's (1980) claim that Wittgenstein's model of similarity was influenced by Peirce's notions of 'continuity'. The essence of Wittgenstein's model is the notion of **projection rules** which are not absolute,

"Things are similar only with reference to a rule of projection. This is often obscured by the fact that we generally use only one rule, that of the portrait which must be similar to its original. But traffic movements are similar to the red, amber, green of the traffic lights."
(1968, p41)

and,

"The shadow is supposed to be **similar** to the fulfillment. But there is no absolute similarity. Similarity varies according to the mode of projection; but we tend to assume, when we use the word, one particular mode of projection, the mode we assume in looking at a photograph."
(p44)

Similarity judgements are thus taken to be made of an **abstract** projection of the entities in terms of their functions. The notion of a 'projection rule' of course directly corresponds to the implicit cognitive operations of the distance models of similarity; and this connection will be developed further in the latter part of this chapter. Another aspect of Wittgenstein's ideas, whose development is also postponed, concerns his discussion of the criterion of a 'good' portrait. He points out that this criterion is **not** similarity per se, but involves intention. The problem remains that there is no simple explanation for intention, "You cannot anticipate the result which you project" (p32).

It is unfortunate that Wittgenstein's innovative ideas have not been taken up in the psychological literature, although the development of our model will partly remedy this omission. Significantly for our argument,

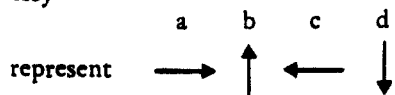
Wittgenstein also addresses the representation problem in relation to the concept of similarity. Playing a piano from a score is cited as an example of understanding a thought in terms of being able to translate it according to some 'general rule' (p44). However Wittgenstein points out that, "- the score does not **cause** us to play as we do; if it did there would be no right and wrong way of playing" (p44). As an example of such a translation process, Wittgenstein defines the letters a, b, c and d as arrows; as in the key to Figure 2.1. The sequence 'a a d a b c' then becomes cell [meaning 1, grammar 1] of the same figure. Wittgenstein points out that "At first sight we would say there was no similarity between the letters and arrows" (p44). The translation rules correspond to the rules of a grammar and, "No description of the world can justify the rules of grammar" (p44).

The complete figure can be seen as an extension of Wittgenstein's argument, with the two entities shown having different meaning and being transformed by two different grammars. The source of the transformation is of course already at an arbitrarily symbolic level, in that letters are symbols of some order. This figure provides the rationale for an experiment in Chapter 4 which entails the encoding of numeric data into alternative symbolic representations, and the evaluation of consequent perceptions of them. The representational implications of these ideas will be developed further in the final chapter. The important point for our argument here, is that judgements of similarity made of the projections resulting from grammar 1 might differ legitimately, in psychological terms, from those made with grammar 2. We will argue that the nature of judgements made concerning the projections made with grammar 1 would correspond to a synthetic cognitive process, and those from grammar 2 to an analytic.

Conclusion

The notion of the similarity relation is of great importance to many philosophical and psychological theories. However the definition and usage of the term both in the literature and everyday life is seldom clear cut or

Key



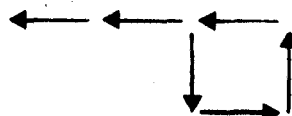
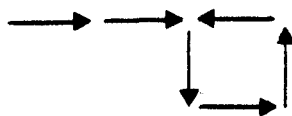
'meaning' 1

a a d a b c

'meaning' 2

d a b c c c

grammar 1



grammar 2

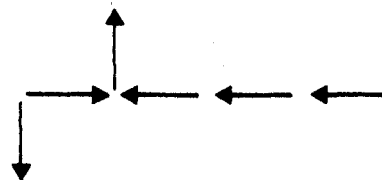
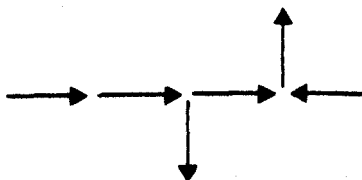


Figure 2.1 Transformations by different grammars

without confusion. The purpose of this review has been twofold: to make clear the origins of various psychological models of similarity that we are to review; and to indicate how the dominance in natural language of the analytic view of similarity relations is likely to have influenced its use in the psychological literature. Gregson, an influential author in the psychological literature, describes similarity as a "pervasive, useful and treacherous idea" (1975). However, he argues that experimental investigation only is useful, and decries the necessity for any philosophical inquiry.

Rather than merely bemoan the arbitrary usage of a complex term in this way, it may be beneficial to examine its origins and use in natural language. A detailed study of the etymology, or even the present use, of the English terms 'similar', 'like' and 'resemblance' is unrewarding. We have already shown that the Sanskrit language has two carefully delineated terms for the analytic and synthetic aspects of similarity, although the difference has apparently been lost in common language. However it is of some interest to note that in both present day German and Arabic, there still exists a distinction between analytic forms of similarity related to notions of simple equality, and synthetic forms concerned with the sameness of function. Further details are contained in Appendix G.

It is not clear whether the distinction that we are claiming is maintained in general speech, although it is certainly not in the German psychometric literature; only the analytic term 'gleich' appears to be used in experimental instructions requiring a similarity response. For the purposes of our argument, 'similarity' will be used as a generic term. Its analytic and synthetic aspects may then be distinguished by the use of the terms 'likeness' and 'resemblance' respectively.

2.4 Psychological models of similarity

2.4.1 Overview

"The notion of similarity is a remarkably pervasive one in psychological theory, of both discursive and formalised sorts. Like most powerful and widespread ideas, it is not amenable to a ready and precise definition -."

Gregson (1975, p3)

A historical and philosophical perspective on the meaning of similarity has been developed, and a distinction derived between analytic and synthetic forms of similarity judgement. An important categorization of models of similarity in the psychological literature was provided by Ekman and Sjöberg (1965). The basis of this dichotomy was a concern for the underlying definition of similarity implicit in the various models. The 'content' models, which would include Tversky's set-theoretic formalism, are taken to assume a heuristic definition of similarity as the quotient of common stimulus content divided by the total content. The epistemological origins of the content models are clearly based on the notion of similarity as a form of partial identity. The 'distance' models are those taken to assume an isomorphism between similarity and distance within some cognitive space. A formal description of such models is that they treat similarity as a monotonically decreasing function in n -space, and yield a representation with no fixed origin. Although the analytic-synthetic distinction may be considered directly with respect to the different cognitive processes that the terms imply, it also has implications for the representational aspects of the objects being considered. It is intended to take Ekman's concern for the model of similarity implicit in the various theories a stage further, and examine the underlying representational implications of the models. We shall therefore use the terms 'content' and 'distance' to describe the models in the psychological literature, but with this widened scope of differentiation.

A central problem for any review of this literature, is that the universal nature of the concept of similarity has caused it to be used and defined differently in many sub-fields. The most important source for this review has been the series of articles on this topic that have been published at regular intervals in the *Psychological Review*, since its inception in the 1890's. These appear to reflect the changes in the disciplines working 'Zeitgeist', the most obvious trend being a continually narrowing perspective on the subject. Recent articles are often little more than detailed expositions of a particular formalism. The desire to present a 'coherent structure', and for 'the clarification of logical status' of the various models of similarity, is reflected in Gregson's (1975) major work in this field, and is our second principal source. We argue that attempts such as Gregson's to distinguish between the various theoretical models have tended to take an insufficiently general initial perspective of the problem. Gregson's 'clarification' for instance is only made within his own formal 'logical' definition of the task.

2.42 The literature

"The question 'What makes things seem alike or seem different' is one so fundamental to psychology that very few psychologists have been naive enough to ask it".
Attneave (1950, p516)

Early approaches

The origins of the conceptual distinction between 'content' and 'distance' models of similarity are clear from our review of the philosophical literature. The writing of the early psychologists also indicated an awareness of this possible distinction in similarity judgements. Herbert Spencer described similarity as the basic relation underlying reasoning and classification,

"Of all relations the most complex is that of Similarity - that in virtue of which we range together objects of the same species, notwithstanding their differences of magnitude, and in virtue of which we group under the same head, phenomena of causation that are widely contrasted in degree." (1881, p256).

He described 'ordinary' similarities between natural objects as being made up of many 'component' or 'simple' similarities; essentially of the attributes of the object. He also writes of 'similarities of sequence' and thus derives the two classes of 'similar coexistences' and 'similar sequences'. Instances given of these classes are 'similar triangles' and 'the uniform sequence of heat upon compression' respectively. By the latter, Spencer meant the similarity between the various cases in which compression produces heat (and was presumably as topical an example 100 years ago as are our present day computer metaphors!),

"Two triangles may be similar, though any side of the one is a score times as great as the homologous side of the other; and though to-day a small disengagement of heat results from the pressure of a hundred pounds, while to-morrow a great disengagement results from the pressure of a hundred tons, the cases are classed as similar." (p258).

From the two classes, Spencer derives, "the two orders of Similarity - perfect and imperfect: the similarity on which mathematical reasoning proceeds and the similarity on which the reasoning of daily life proceeds." (p258). To Spencer, 'perfect' similarity is concerned with the matching of relations whilst, "In imperfect similarity, however, the only specific implication is - likeness of nature in the relations." (p259). It would seem legitimate to describe this as a synthetic similarity judgement, involving the process of 'perception in context'.

William James was more explicit in his delineation of the problem. The quotation in section 1.1 concerns his comparison of a theory of 'immediate resemblance', which he attributed to Stumpf, to a model of likeness 'dependent on partially identical content'. The importance of the treatment of the perceptual component of any general model of similarity 'judgement' will be argued later. James' model would appear to have been strongly influenced by Herbart's notions of 'schematic implicit apperception'. Describing a botanist examining a flower, Herbart's (translated in McDougall, 1923) writing suggests a 'top-down' model,

"The whole system is excited from the first; and the systematic relation of the parts governs the order of the perceptual activity. The parts may be said to be implicitly apprehended throughout the process, while each part in turn becomes explicitly apprehended." (p259)

Gregson (1975) dismisses James' argument as 'facile', arguing that because simple impressions are undecomposable James' 'immediate resemble' or similarity measure is therefore unanalysable. Despite providing a theory that does not fit into Gregson's formal notation scheme, James' incisive criticism of the partial identity or content models is equally relevant ninety years later in relation to the present 'contrast' models of similarity; with their set theoretic basis. The argument he puts forward stems from the philosophical distinction between the 'essential' and the 'accidental' features of an object. In the real world, this of course depends upon the context of the object and the perceptions of the individual rather than the arbitrary decision of the experimenter. James introduces the intriguing term 'Mind-dust theory' to describe such models, presumably in contrast to associationistic ideas.

Kulpe is critical of the simple associationist view that "-contiguity is the sole incentive to association, and that all apparent cases of reproduction by similarity or contrast must really be referred to it." (1895, p192). He provides an equivalent but more formal classification of similarity types

which would also indicate an awareness of the content-distance distinction,

"The similarity of two simple qualities may consist (a) in the slightness of the difference that exists between them - (b) Or similarity may be defined as partial identity. - (c) Or, lastly, similarity may be predicated of two qualities which stand in one and the same relation to a third. - Only the second of the three possible meanings of similarity can be at all precisely formulated." (p192)

Kulpe gives as an example of his first definition, "two just discriminable shades of indigo blue". He would almost certainly have conceived of it in distance terms, although present day psychometricians might formulate it in set-theoretic terms. The second definition clearly describes a 'content' model, he refers to the partial match of colour tones in terms of their attributes of saturation, extension, duration and quality. Kulpe's third definition resembles the formal description of a distance model, and Gregson appears to treat it as such. However from the concrete example that he provides it is not clear what was meant, "Thus red and green are similar, because both reproduce the word 'colour'."

Ink blots and handwriting

Dearborn (1910) attempted an analysis of the mental processes underlying similarity and dissimilarity, the discernment of likeness and unlikeness. One hundred inkblot cards were arranged in a 10 x 10 matrix and the subjects were required to pick out the ten most similar, or dissimilar, to a norm card. It is noteworthy that this procedure is conceptually very similar to the 'reference ranking' task that will later be discussed in the experimental section, in that it permits the analysis of the subjects unconstrained ranking procedure. Dearborn's choice of ink blots as stimuli unfortunately presented problems in terms of the formal rigour with which any analysis of the subjects' judgements might have been made. Also it would have been preferable if he had relied on the assessments of several judges

rather than, as he apparently did, just on his own. However in relation to the range of stimuli currently being employed in experimental studies it is interesting to note that he writes,

"Only very rarely, moreover, would the actual objective similarity in life be as narrowly confined as in the conditions of these simple experiments" (p61).

Dearborn reported that only one subject, a music student, claimed to have any true 'feeling' of likeness or of unlikeness, and this subject's judgements were considered the best. Some subjects attended only to the actual visual sensation, whilst others attended to specific characteristics of them. Dearborn describes this as 'looking at' versus 'thinking about' the blots. In the sense of our argument, Dearborn explains the more 'satisfactory' choices of the student who experienced some 'feeling' for the blots, and also those who 'looked at' them, in terms of their attributes and relations,

"Even these simple bluish-black forms in only two dimensions have so many characters that to specify one, or two, or three, and compare them by these leads to imperfect and misleading results. On the other hand, the "feeling" of likeness or of unlikeness implies a much wider acquaintance with the blots and is, therefore, the basis of a better comparison." (p60)

Following Dearborn's study, Hollingworth made a study of similarity and dissimilarity judgements between different samples of handwriting. His results also suggested the existence of different judgement processes (as had previous 'handwriting' studies in France by Binet, and other researchers in the U.S.A.),

"Judgements within each type or category involve each its own peculiar psychological processes and criteria. The 'most similar' is not by virtue of that fact, the 'least different' -" (1913, p289).

Hollingworth made correlational analyses of the accuracy of similarity and dissimilarity judgements over repeated trials, and found similarity judgements to be best. Hollingworth's main claim from an analysis of subjects protocols is, "the greater ease and naturalness which is felt to characterize the judgements of similarity." (p287). His subjects found it difficult to define adequate criteria for detecting differences, and also in using any such analytic criteria consistently, "The judgement of difference - is largely or often based on the comparison of fine points and minor details." (p287).

Hollingworth suggests that similarity is a more fundamental judgement than dissimilarity, "- in real life, it is similarity that most interests us." (p275). He takes issue with Dearborn who he claims had put forward the idea that dissimilarity is the more fundamental relationship. Gregson too misunderstands Dearborn on this point, misquoting his argument. Dearborn distinguishes between judgements involving 'feeling for a wider acquaintance' and the 'specification and comparison of characters'. His actual claim is that only in the latter case, the analytic matching of attributes, are dissimilarity judgements easier to make,

"Some ideal criteria were obviously more essential than others and led to the selection of a set of blots evidently like each other and the norm. **Ideal criteria gave more accurate results in the dissimilarity choices than in the similarity choices. This is, as we should expect, on logical principles. The awareness of unlikeness is an easier, if not a simpler, process apparently than that of likeness, for the change of consciousness is greater and so easier to appreciate. At any rate, the sets of blots chosen as unlike the norm were much more certainly unlike it than were the "similar" blots chosen like it.**" (1910, p61, Gregson's quotation emboldened)

Given his previous comparison between judgements involving 'feeling' as against articulated characterizations, the meaning of Dearborn's phrase 'more easily appreciated change of consciousness' is quite clear.

Hollingworth does not seem to consider that unlike his own subjects, many of Dearborn's were able to differentiate characteristic attributes of the stimuli. An interpretation of Hollingworth's results in terms of the framework being put forward here, is that it is clearly easier to make synthetic similarity judgements than analytic dissimilarity ones.

Gregson appears to willfully misunderstand Dearborn's argument, perhaps because the conceptual basis of his own analytic framework assumes an underlying unitary process. In misquoting Dearborn as we have noted, he appears to be attempting to create the situation where he may then dismiss all the important issues raised in the early literature either because of formal inadequacy or disagreements, "The assertion that likeness is more basic than unlikeness has been flatly contradicted by other writers of the time with even less evidence." (1975, p33). We have already remarked on Gregson's unwillingness to examine more than superficially any philosophical issues. His cavalier attitude to the early literature is also clear, "One is hardly obliged to subscribe to the views of any of the earlier writers as almost no experiments on similarity with a coherent structure were reported by them." (p27). What Gregson does not appear to consider is that the early psychologists may have been more aware of, and more willing to consider, the underlying conceptual problems. It is now difficult to conceive of a Psychological Review article examining different philosophical doctrines in an attempt to understand the nature of perceived relations in similarity judgements, as did Dunlap's in 1912. We shall later note that modern psychological texts (e.g. Miller and Johnson-Laird, 1976), do little more than 'hand wave' at such seemingly intractable problems as the correspondence between the relations and the attributes of an object. In the present literature, the deeper problems concerned with judgements of similarity have effectively been trivialized to the form of simple mathematical matching operations, in order that experimentation of a seemly rigour may be undertaken.

Later Approaches

One of the most thoughtful and influential reviews of this topic has been provided by Attneave (1950). He indicates the importance of the concept of similarity to major areas of psychological inquiry and suggested that a 'comprehensive analysis' was necessary. Although individual studies might 'attack no more than a small segment', Attneave considered it unfruitful to adopt arbitrary operational definitions or various forms of 'similarity' for the various sub-problems. It is unfortunate that Gregson does not discuss this work in any detail, because Attneave's treatment of similarity models is perhaps the most clear in the whole literature,

"It is obvious that when things are similar they are similar with **respect to something**. The characteristics with respect to which objects are similar may be conceptualised either as more or less discrete and common elements or as dimensions on which the objects have some proximity." (Attneave, 1950, p519)

The main part of Attneave's paper is concerned with a discussion of the problems of representing a possibly complex 'psychological space' within a simple physical Euclidean space. Further detailed discussion of this study will be made in a later section:

Wallach's (1958) review paper was the last in the literature to seriously examine the philosophical and representational problems of models of similarity. He delineates three definitions from the literature. The first is traced back to Hume's discussion of similarity as partial identity. Similarity is consequently defined in terms of **common environmental properties**. An alternative approach noted by Wallach is to "reorganize the organism's selective functions" (p105). This leads to a definition of psychological similarity as the **making of a common response to common environmental properties**. He shows that the third view, of similarity as a **primary stimulation gradient** follows from associationist conceptions.

Neural traces are thought to be laid down along various dimensions when a stimulus impinges, and 'psychological similarity' depends on how far a new stimulus is from an old stimulus on such a dimension. This is effectively a physiological exposition of a cognitive 'space' model. Wallach also develops a fourth model of his own, which is really only a more complex version of the second. He suggests that "recognition of similarity depends on applying a rule which leads one to **assign items to a common category**" (p106). He likens the idea to Kohler's term **distinctive similarity**, which denotes particular items that resemble each other more than they resemble the rest of the situation. This is one of the few papers in the literature of which Gregson is not critical and appears to form the conceptual basis of his classification scheme.

Following Attneave and Wallach, the only attempt to take a general perspective on this topic has been the book by Gregson (1975). The most common role of similarity in the psychological literature has been its wide use as an explanatory principle; e.g. in association theories of memory and transfer phenomena such as stimulus generalisation and retroactive inhibition. It is in this historically primary role that imprecisely defined measures of similarity measurement have caused the greatest problem. Attneave pointed out the "confusion and ruin" that had come to psychological theories built on such imprecisely defined similarity measurement. Gregson's detailed work also indicates how ill founded many of these ideas are, and he puts forward various categorisations in an attempt to overcome the problems. He proposes a formalism for describing similarity judgements in terms of 'levels' of relative judgement. His categorisation scheme appears to rely heavily on 'information' and 'set' theoretic axioms; the first two major assumptions being,

"1.31 Similarity is a function from pairs of stimuli onto the closed real interval $[0,1]$, which empirically we will treat as a continuous variable in the context of models of behaviour.

1.32 A similarity response is made to a pairwise comparison of two stimuli, each of which is a complex presentation of the form Signal with Noise." (1975, p16)

Another important assumption underlying Gregson's formalism is a fundamental split between perception and judgement. Within the process of a subject 'making a similarity response' he considers that errors might occur "either because the stimulus properties are incorrectly perceived or because his similarity judgement process is in some way inappropriate" (1975, p16). However his treatment is concerned only with judgement processes. Although Gregson makes a wide ranging review of literature concerned with similarity, as we have remarked earlier, he is dismissive of attempts to examine the philosophical and general use of the term. Indeed it is difficult to see how he might have made use of any such possible insights within the strictly formal model that he develops. We shall later note that Gregson's work clearly falls within the 'purist' psychophysical tradition. Gregson's most original contribution to the subject would appear to be the derivation of a formal notation for classifying the level of complexity of relative judgements,

L⁰: absolute judgements; a stimulus compared with the null stimulus noise

L¹: pairwise comparison judgements i.e. between two stimuli.

L²: relative pairwise comparison judgements i.e. a judgement made between a pair of pairs of stimuli.

L³: etc

The following operators are defined within Gregson's classification system,

S_j is the jth signal

N is the noise

+ is a concatenation operator and thus the jth stimulus is S_j + N

./ represents "binary relational judgements operations"

./. represents an operation of relative comparison

The levels of judgement then become,

$$L^{\circ} = S + N./N$$

$$L' = S1 + N./S2 + N$$

$$L^2 = S1 + N./S2 + N./S3 + N./S4 + N$$

etc

Gregson's theory is based on the idea that judgements directly map the physical correlates of stimuli, the fundamental operation is therefore of discrimination at L° . Gregson contrasts his model with that of Krantz (1972) who had earlier developed a 'relational theory' in which a judgement equivalent to L' is taken as primitive instead of L° . Gregson also refers to another well known classification of psychological data, Coombs' (1964) 'Theory of Data'. Coombs' model of similarity will later be developed in more detail, but essentially his classification system is based on a distinction between single stimulus or event situations and dyadic situations. Similarities data is here defined as a relation observed, "on a pair of distances where each distance is between a pair of stimuli". Gregson claims this definition to be more "restricted" than his own and writes that,

"It is clear that in Coombs' system similarities data is L^2 with $./$ replaced by a psychological distance measure and $./.$ replaced by a distance comparison operation $-$." (1975, p21).

Gregson's formal axioms require the fundamental relational function to be a binary operator between pairs of stimuli. This appears to be quite coherent in terms of 'discrimination' at L° , but causes great problems when he comes to discuss in detail the more complex forms of similarity judgement tasks. In a section of his book concerned with problems in experimental studies, Gregson refers to similarities data derived from triadic comparisons to be, " L' over pairs within triples" (1975, p104). The literary theories that he so decries then become quite clear in contrast to his consequent discussion of the "suppressed similarities" between

non-referent items, and the necessity to conceptualise levels " intermediate between L^2 and L^3 " (1975, p161); which are of course logically impossible within his formalism. It is clear that Coombs' similarity model cannot be described in several fundamental respects by Gregson's L^2 , and should not be thought of as being 'more restricted'. In the same section concerned with experimental studies, Gregson writes of the many reported studies that are imprecise in stating what subjects are actually asked to do. Gregson notes that studies "may be ostensibly at L^1 in terms of instructions, but in terms of the task given at L^2 , or the other way round." (1975, p159). This appears to be precisely the error that he makes himself when attempting to define his various levels. His attempted description of a possible L^1 experimental situation is itself quite confused as to whether the process is one of discrimination or similarity judgement, and the subjects perception of the task that he describes is clearly at L^2 (1975, p18).

Whilst he recognises that 'context effects' are likely to occur with experiments at L^2 , Gregson's formal system is not sufficiently general to deal with them. He notes that such considerations of context and strategy create difficulties, "which are enhanced as the complexity of experiments increases", and is forced to conclude that "it can be easier to analyse judgements with a model at L^2 , permitting L^1 as a degenerate case, rather than the reverse" (1975, p160). Furthermore Gregson admits that subjects may be asked to judge on either "undefined overall similarity" or "with regard to specific dimensions of variation". Whilst noting these important effects, Gregson again cannot address them properly as they are, "independent of the distinction between L^1 and L^2 " (1975, p159).

The principle concern of our argument is of course to deal with just such problems; in particular the individual's use of 'overall' or 'specific' judgement strategies in their perceptions of similarity. Accepting L^1 as being a 'degenerate' case in terms of experimental specification, also does not support Gregson's earlier claim for greater generality over Coombs' model of similarities data.

The main reason for the influence of Gregson's work, and our detailed treatment of it here, is that he was the first to propose that the 'content-distance' dichotomy is an unsatisfactory classification for models of similarity. It has already been made clear that the content-distance distinction is an important aspect of our argument in terms of contrasting the existing models in the literature, and putting forward improvements. Gregson (1976), and later Junge (1978), claimed that it is possible to distinguish a class of psychometric models, the 'normed distance' or 'point projection' models, which can be written equally well in 'content' or 'distance' forms. Gregson has argued that Eisler's definition is not sufficiently restrictive in that he has shown "a wide diversity of ways of defining common and total content" (1976, p143). This is not difficult, because in representational terms it is of course quite arbitrary for the systems of binary attributes that Gregson wishes to assume. The principle argument against Gregson's claim is that it is not necessarily psychologically significant that, under certain constraints, two models can be shown to be mathematically isomorphic. For example, Roskam (1979, p224) shows that Tversky's set theoretic 'contrast' model of similarity can be shown to have the formal 'distance' properties of the 'city-block' metric; a dimensional model. Yet the psychological assumptions underlying this model as to what constitutes a judgement, what similarity means, and the implicit representation of cognitive objects are quite different to those made by the 'distance' models of similarity.

For further evidence that Gregson's claim is not well founded in relation to general models of similarity, it is necessary to examine some of the basic assumptions of the 'psychophysical' and the 'scaling' approaches to psychometrics. In the scaling approach the investigator **assumes he knows the appropriate spatial model**, or that the particular spatial model will not grossly distort the assumed cognitive representation. The judgements of similarity are treated as distances in the assumed n -dimensional psychological space, and a possible specification is extracted in some sufficiently low dimensionality. Authors such as Gregson, and also Beals et

al (1968) are critical of the scaling approach which is "committed a priori to spatial representations of hypothetical processes underlying relational judgements." (Gregson 1976, p140). However this is not a substantial argument in that 'data reduction' as a form of analysis is not necessarily atheoretic. We shall discuss later Takane's work, using the recently developed maximum-likelihood scaling algorithms to test alternative hypotheses of both representational structure and judgemental process.

In the psychophysical approach the investigator **assumes he knows the component dimensions or attributes,**

"- the physical specification of a set of M stimuli is known and each X in M can be written as a vector of measures of components x_1, x_2, \dots, x_t the components or dimensions being numbered 1 to t . The associated $M \times M$ similarity measures, each of which is $P(X, Y)$, a number between zero and one, are known empirically and can be predicted, theoretically, from some model C_s which takes as arguments the physical specifications of the stimuli."
Gregson (1976, p140)

The investigator then attempts to directly decide which combinatorial rule best describes how the subject uses these 'known' dimensions or attributes in making an overall judgement. The psychophysical approach is comparative in that the data is used to decide between alternative models for describing judgements. The term 'model' being regarded simply as some form of parameterized combination rule. Thus the core a priori assumption of the psychophysical approach is that it is possible for the experimenter to provide an objective description of all possible aspects of the subject's perception of the stimuli. It is from this very limited perspective of knowledge representation that Gregson extols the 'formal' over the 'literary' models of similarity judgement. In comparing the two approaches, it might well be argued that the a priori restriction of possible knowledge systems, as against the a priori assumption of spatial representation, is the more serious limitation. Therefore with regard to general models of

similarity we shall not accept Gregson's argument that, "The simple content distance dichotomy is a historical mistake, best discarded" (1976, p143), but instead continue to regard it as an important criteria for determining the interrelations of the various models in the literature.

2.43 Distance models of similarity

Introduction

Due to their concern for what can be regarded as synthetic judgements of similarity, it is clear that many of the models in the early psychological literature may be characterized as distance models. With regard to the more recent literature, it is impossible to separate the development of distance models from the development of the multidimensional scaling (MDS) analyses of multivariate data. The essence of MDS analysis is a distance model of similarity,

"We conceive, then, of a psychological space with one set of points in it corresponding to the set of stimuli, and with the relations among these points to correspond, in some sense, to the relations among the stimuli as given by behaviour. In particular we conceive of a distance between every pair of points in the space which is a measure of the **similarity** of the corresponding stimuli." Coombs (1964, p433)

Critics of MDS have often rightly pointed out the sometimes circular relation between MDS **analysis** and cognitive structure models. It is important to note that much of the literature concerned with MDS analysis is not primarily concerned with psychological models of similarity. Many of the scaling algorithms were effectively developed as forms of data reduction, or for the analysis of the 'pure' scales of psychophysical phenomena. The underlying distance model of similarity received little consideration. We shall later note that psychometricians are becoming more concerned with the

actual processes of data generation, and with the development of adequate models for the analysis of more complex forms of data. The main part of this section will be concerned with detailing the MDS models. However it will be useful to first broaden the discussion and examine the general use of notions of cognitive 'space' and 'distance' in the psychological literature.

Cognitive space and the metaphoric use of 'distance'

The first explicit use of the concept of 'psychological distance' can be traced back to Isaac Newton's use of the colour 'circle' in 1704, and the German psychologists of the early nineteenth century probably developed their use of the notion as a consequence of their studies of perceptual phenomena. However, Bertalanffy (1973) points out that the Indo-European languages have often employed spatial metaphors to express nonspatial relationships: long and short for durations; heavy, light, high and low for intensity; and approach, rise and fall for tendencies. He quotes Lorenz as having pointed out that the concept of time is also represented in terms of a visualisable model of space. Spatio-temporal processes are implied in expressions such as 'the course of time', our use of the prepositions 'before' and 'after' and the nouns 'past', 'present' and 'future'. Lorenz also apparently proposed the intriguing notion that the tactile space still has a particular preponderance over the visual. We have 'grasped' a 'connection' only if we can 'comprehend' (*begreifen*, i.e., seize) it. Even the idea of object, (*Gegenstand*, that which stands against us) originated in the haptic perception of space.

Some theorists in the psychological literature make explicit formal claims regarding the nature of psychological 'space', whilst others either implicitly or explicitly make use of it in a purely metaphoric role. For example the 'dichotomy corollary' of Kelly's (1955) personal construct theory assumes the individual's 'construction system' to be composed of a finite number of dichotomous constructs; with an organisational structure of constructs. Kelly claims that this corollary assumes a structure of

psychological processes which lends itself to mathematical analysis, although it is not clear whether this circumstance is purely fortuitous. The assumption of the dichotomy corollary itself is also suspect, in that the Heraclitean principle of thinking in opposites has long been regarded as a limiting feature of many languages (Ogden, 1967).

There is a general problem, in that theorists tend to describe psychological space purely in terms of their own analytic technique, even though it appears fatuous to consider that any necessarily limited set of dimensions, classes, chunks, etc might provide more than a partial representation of knowledge. In some models, the mistake is also made of confusing an operational definition of the measure of cognitive space with the definition of the concept itself. Gregson (1975) points out just such a "classic mistake" in the Semantic Differential theory of Osgood, Succi and Tannenbaum. They do not explicitly define the meaning of similarity but instead write, "In this representation we can 'see' the similarity between various concepts on all factors simultaneously in terms of their closeness in the space." (1957, p89). Gregson's argument is that they have taken a spatial model as axiomatic, and superimposed a metapsychological explanation after the factor analyses had been found to 'work', "The meaning of a concept to an individual subject is defined operationally as the set of factor scores in the column representing that concept." (1957, p87).

It is interesting to note that several publications have explicitly addressed the issue of the spatial metaphor relating the concepts of similarity and distance. Cunningham and Shepard (1974) identify two distinct objectives in the attempts to formalize the implicit connection between similarity and distance. The central objective in applications of MDS has been to infer the structure of the underlying space. The second, which is Cunningham's concern, is the determination of the shape of the underlying psychophysical function. To this end he develops a method of analysis which he terms 'non-dimensional' scaling. This form of analysis may be used to determine whether a particular data set will be least distorted, or best

represented, by the assumptions of the hierarchical or scaling analysis algorithms. Shepard and Arabie (1979) developed a nonhierarchical additive clustering model which is explicitly not based on any notion of distance. Unfortunately in order to 'escape' the distance metaphor they have to posit a rather simple representational system of discrete binary valued properties.

Illustrative of the communicative power of the distance metaphor is the fact that even those theorists who are most critical of the acceptance of the metric axioms, nevertheless treat similarity as a complement of distance. Thus Tversky writes "Similarity is a relation of **proximity** that holds between two objects." (1977, p347, my emphasis). Psychologists have often attempted to incorporate novel mathematical formalisms into their theories, in the form of explanatory metaphor e.g. the use of field-theory and fuzzy set theory. An illusionary enlightenment often follows from the purely metaphoric use of notions of cognitive space in the psychological literature. The power of the metaphor can give an apparent sense to an inherently unsound model. Several of the more reflective researchers concerned with the application of multi-dimensional scaling models as a cognitive representation have expressed concern about this issue. In a paper primarily devoted to questioning the formal existence of the factor model, despite its use for 50 years, Schonemann ends,

"it would be unfortunate if, a few years hence, someone were to discover that multidimensional scaling does not exist because the concept of a distance was not understood well enough." (1977, p812).

Multidimensional scaling

Multidimensional scaling (MDS) is a powerful mathematical procedure which is used to systemize data. MDS procedures represent objects which have been judged to be similar to one another as points close to each other in a spatial format. The problem that they have generally been directed at is to

measure and understand the relationship between objects when the underlying dimensions are not known. The first scaling analysis was made by Richardson (1938), using psychophysical data. These 'metric' scaling procedures were later further developed by Torgerson (1958), but made very strong formal requirements of the data. The perceptual or cognitive structure of the set of stimuli had to be Euclidean in nature and the observations of similarity of pairs of stimuli had to be linearly related to distances between points in space. The next development in scaling analysis was the nonmetric 'unfolding' approach originally developed by Coombs (1964). This procedure was most useful for describing the interrelations of the judges and provided little information concerning the spatial configuration of the objects. Ekman (1954) had also suggested a vector model for MDS, where similarity was interpreted directly as a scalar product angle, an approach "that seemed to give complex solutions for rather simple data." (Torgerson, 1965, p380).

The most significant advance came with the development by Shepard (1962) of an 'interactive' algorithm for finding the appropriate monotonic transform of the original data to the 'best' fitted dimensional structure. The nonmetric procedures make the weaker assumption concerning similarity, "- that it be the complement of distance in a space of one kind or another." (Torgerson, 1965, p379). This implies a monotonic decreasing function f , so that

$$s(x,y) = f[d(x,y)] \quad 2.43.1$$

where $s(x,y)$ denotes the observed similarity between x and y , and $d(x,y)$ denotes the distance between the corresponding points in the stimulus configuration. The generalised distance function is given by,

$$d_{xy} = \left(\sum_{k=1}^K (n_{xk} - n_{yk})^2 \right)^{1/r} \quad 2.43.2$$

where the distance between points is specified in terms of K -dimensional

space of Minkowski r -metric. For $r=2$ this represents ordinary Euclidean distances. For $r=1$ the distances are known as 'city-block', and for $r=\text{infinity}$ the formula gives the maximum of the absolute coordinates, the 'dominance' metric. The non-Euclidean cognitive 'space' implied by the 'city-block' and 'dominance' metrics is more difficult to conceptualise. Attneave (1950) and Torgerson (1965) have shown that these metrics are most suited to occasions where subjects are consciously aware of the dimensions that they are manipulating, and in this sense, they can be considered to be degenerate spatial models. Arnold (1971) showed that the simple dominance metric provides a better account of data collected by means α procedures which impose severe information-processing constraints on the subject.

The importance of the data-collection process in this context is pointed out by Coxon and Jones, "it suggests that methods which are more attuned to the analytic rather than the synthesising aspects of information-processing will differentiate and break up the apparently refractory nature of sociological concepts such as 'prestige'." (1979, p39) In an introductory text concerned with the use of MDS analyses, Schiffman et al (1981) write, "For various technical reasons, however, we generally recommend that dissimilarity data be gathered, and that the user convert similarity data into dissimilarity data in whatever way seems reasonable." (p77). These authors would appear to be biased by the form of stimuli with which they are most concerned, the taste and odours of food and drink products. The advice is likely to have unfortunate consequences, except where the stimuli are 'simple'; i.e. they really can be described in two or three dimensions. For more complex stimuli, the most probable outcome is for subjects to be directed towards, and adopt, an analytic decompositional approach to the task. Our argument is that they may have been thus directed towards an inappropriate strategy, even though the results may appear to be more consistent.

The operation of the scaling algorithms is essentially one of statistical fitting. A configuration is found, by an iterative process, whose distances best fit the data. The process being controlled by the minimisation of

some general loss function. Kruskal (1964) was the first to formally specify such a 'badness of fit' or 'stress' formulae,

$$S = \frac{\sum_{x,y} (d_{xy} - \hat{d}_{xy})^2}{\sum_{x,y} d_{xy}^2} \quad 2.43.3$$

Kruskal introduced new fitting quantities $\hat{d}(x,y)$, which are referred to in the literature as 'pseudo distances', 'disparities' or 'discrepancies'. They are the best squares fit to the distances $d(x,y)$ and are as close as possible to having the same ordinal structure as the data. There has been a continuous development of the MDS analyses, providing not only more powerful algorithms than can cope with 'noisy' data, but which also suggest greater insight as to the possible underlying judgemental processes. Following Tucker (1960), various methods have been developed for bringing out individual differences in judgement. In the INDSCAL model of Carroll and Chang (1970), individual differences are parameterised by individual sets of 'importance weights' for the dimensions of a common space. The PINDIS model of Lingoes and Borg (1978) makes fewer restrictions upon the position and relative rotation of the axis of each individual's 'space', but can treat only metric data. Ramsay (1977) and also Takane (1981) have developed algorithms using maximum likelihood estimation techniques, which enable the analyst to test both hypothetical data structures and hypothetical judgement processes. De Leeuw and Heiser (1980) have also recently detailed a method of 'constrained-MDS' which, given a testable hypothesis, is capable of differentiating small inter-group differences in fairly 'noisy' data.

The implicit models of similarity in MDS

Having charted the development of MDS as an analytic technique, we can now examine it in terms of the changing conceptualizations of the underlying distance models of similarity. We have already referenced Schonemann et al's (1979) discouraging analysis of the **exploratory** use of MDS as

"mechanical data reduction". The authors of the new confirmatory scaling methods (e.g. Lingoes and Borg 1978) would appear to be actively attempting to discourage such "blind" misuses. Roskam (1979) provides an account of the implications of the axiomatic propositions underlying data theory and mathematical models. Roskam argues that multidimensional scaling is a form of 'derived measurement'. He points out that there are no rules for choosing the right model to analyse ones data and that,

"the domain of validity of a model has to be determined from empirical testing, that is: by **goodness of fit** (which expresses the risk we are willing to take) and by the criteria of **plausibility** and **interpretability** of results." (p228).

Both the proponents and opponents of the spatial models have tended to assume that the correspondence with distance must imply continuity, that dimensions and directions within a space must be interpreted as **continuous** variables. However several distance models exist which allow more complex models of similarity to be examined. The idea of unfillable 'holes' in space is not intuitively obvious, but Torgerson (1965) first showed how relatively simple structures underlying a set of stimuli might cause a 'U' shaped surface containing a space that cannot be 'filled'. Torgerson has proposed an alternative structural analysis to that of MDS, factor analysis, and cluster analysis.

"In the ideal type model, the spatial location of a given stimulus point is given by its distance from each of a number of ideal types. Thus the degree of dissimilarity, i.e. the distance, between a pair of stimuli depends on how close each is to the various ideals. The rationale of the ideal type model leads to representation of the stimuli as points in a multidimensional hyperspherical space, where observed dissimilarities are interpreted as angular distances in the positive orthant on the surface of a hypersphere." (1983, p2).

Theoretically the 'ideal type' model is related to both MDS and factor analysis, and also less directly to cluster analysis. However, conceptually the form of judgement that is implied clearly distinguishes it as a distance model. Within Torgerson's 'ideal type' model, the similarity between objects is no longer described by a simple monotonically decreasing distance function. The distance between an object and its nearest 'ideal type' is effectively weighted more heavily. The 'hyperspace' mathematics underlying this model are fairly involved, but Torgerson (1982) has now developed an algorithm which implements the 'ideal type' model. This form of analysis would be well suited to the analysis of data containing 'prototypes' which cannot be adequately described by either the simple distance or the content models.

Following Tversky's (1977) criticism of the scaling models of similarity, Krumhansl proposed a spatial model with a modified distance function, "- the similarity between objects is a function not only of interpoint distance in a metric space but also the spatial density of points in the surrounding configuration." (1978, p446) The 'distance-density' model effectively adds parameters to the standard distance equation. These might be derived from prior experimentation or hypothesis,

$$\bar{d}(x,y) = d(x,y) + \alpha\delta(x) + \beta\delta(y) \quad 2.43.4$$

where $\delta(x)$ and $\delta(y)$ are measures of spatial density in the neighbourhoods of x and y , and α and β are constraints that reflect the relative weight given to densities $\delta(x)$ and $\delta(y)$.

Tversky's criticism of the geometric approaches to similarity has already been noted, and will be further detailed in the next section. However Krumhansl effectively annuls the feature theorists' criticisms of the dimensional model, by proving that spatial density effects are equally well able to explain data 'distortions' such as distinctiveness and asymmetry. She does not directly reference them, but Lingoes (1978) and others have developed, and implemented as programs, a whole series of 'regional' analysis models. These are capable of 'neighbourhood preserving' transformations and thus incorporate spatial density effects. Roskam (1968)

provides a detailed theoretical basis for 'distance-density' models. His compensatory distance model implies some of the properties of the distance models, but there are "subtle" differences in terms of "the uniqueness specified in terms of the boundedness property" (p672). Roskam's formal term, a 'boundary hyperplane subspace', describes a possible model of intransitive similarity judgements.

Takane and Carroll's (1981) paper exemplifies the more recent 'parametric approach' to nonmetric scaling, and provides a more detailed testable model of similarity. Nonmetric data are viewed as **incomplete** data and are considered to convey only ordinal information about distances. An **unobserved metric process** conveying complete information about distances is assumed to underlie the nonmetric data that is elicited from the subject. Takane's algorithms permit the user to test different models of error that might occur in this hypothesised mapping of the underlying process. It is therefore possible to examine the effects of different task and context constraints upon similarity judgements; an important development.

2.44 Content models of similarity

The origins

The philosophical origins of the concept of similarity implicit in what Ekman has termed the 'content' models are clear from our earlier review. It was also noted that the dominant use of the word in present (English) speech is to describe some form of attribute matching process. Gregson traces the first formal description of similarity in terms of set theoretic measures to the 'eccentric' English philosopher psychologist Smee. His treatise of 1851 was entitled 'The process of thought adapted to words and language together with a description of the relational and differential machine'. This machine, which would appear to have anticipated the coming 'fifth-generation' of computers, was intended to operate a rudimentary

propositional calculus. Smee described his machine as being "based on the laws of thought", and using a Boolean like algebra, it was intended to convert assertions of similarity into hierarchical classification schemes.

The reinvention of Smee's ideas in the psychological literature can be traced back to the models of statistical learning, and fundamental measurement theory of the 1950's. The most important source is the paper by Bush and Mosteller which attempted to describe the process of stimulus generalisation in terms of "a model based upon elementary concepts of mathematical set theory" (1951, p413). This work is cited as a source by both Attneave (1950) and Tversky (1977); although it is unfortunately not discussed by Gregson. Bush and Mosteller define a similarity index η ($= n$) as the ratio of the measure of the intersection to the measure of one of the sets. Their definition amounts to an operational evaluation of n ,

"Although there are several intuitive notions as to what is meant by similarity, one usually means the properties which give rise to generalisation. We see no alternative to using the amount of generalisation as an operational definition of degree of similarity."

(1951, p413)

Bush and Mosteller consider the case of similarity between two non-disjunct sets S and S' ,

$$\eta(S \text{ to } S') = \frac{m(S \cap S')}{m(S)} \quad 2.44.1$$

where $\eta(S \text{ to } S')$ defines an index of similarity of S to S' ,
 $m()$ denotes the measure of any set or sub-set
and $S \cap S'$ denotes the intersection of S and S' .

$$\eta(S' \text{ to } S) = \frac{m(S \cap S')}{m(S')} \quad 2.44.2$$

They make clear the tacit assumption "that the measure of an element or set of elements is independent of the set in which it is measured" (p416). The

definition of 'measure' is important, in that other than within the animal-learning paradigm it must signify 'meaning' in some sense. The stimulus situation is denoted "by a set of stimuli which is part of the entire universe of stimuli" (p414). The properties and number of elements are explicitly left undefined. The 'measure' of a finite set is then the sum of 'weights' associated with each element, "Intuitively, the weight associated with an element is the measure of the potential importance of that element in influencing the organism's behaviour" (p414). Bush and Mosteller's index of n thus takes account of variation in the relevance of common elements, as well as the contribution of unique elements. It seemed 'regrettable' to Bush and Mosteller that their definition of similarity is non symmetric by definition,

"However we do not care to make the general assumption that (a) the measures of all situations are equal and at the same time make the assumption that (b) measures of an element or set of elements is the same in each situation in which it appears." (p416).

They cite the example of the 'set of elements' being a light bulb in a small box or in a ballroom. In our later discussion of Tversky's 'context' effect experiments we shall attempt to distinguish between 'local' and 'global' context effects. The former can be defined purely in terms of the attributes of the set of stimuli, whilst the latter is conceived of as the reference base from which subjects try to make sense of the experiment and its stimuli. It is clear that Bush and Mosteller were aware of the possible significance of such effects and also of the deeper representational implications of 'defining' complete sets of elements, "Further this pair of assumptions, (a) and (b), leads to conceptual difficulties." (p416).

There have been numerous amendments suggested to Bush and Mosteller's basic formulation e.g. Noble (1957), Restle (1959). However all of these, including Tversky's (1977) contrast model, can be considered to be generalisations rather than conceptual improvements. Noble criticised the

original model of Bush and Mosteller for not being theoretically coordinated with any stimulus dimensions. His was the first content model to discuss cognitive objects as 'universes of elements' i.e. as an unordered set of elements. Unlike later authors, Noble at least tries to make some sense of such elemental decomposition,

"Some will object to this tacit endorsement of the doctrine of identical elements. - Unfortunately no rules can be offered for the definition of elements. Like the matter of deciding what a stimulus is, or a trial, or what events are reinforcing, psychologists must resort to the usual analytic and pragmatic evaluative procedures which are common to all natural-science undertakings. - Once the elements have been reliably identified, however, the problem of equality of elements may be solved by randomization or counterbalancing techniques within the experimental design." (1957, p35)

Restle first noted that the set theoretic notions allowed a description of similarity that might be conceptualised other than in 'distance' terms; he developed the implications of 'ordering' and 'betweenness' of sets. However, noting Attneave's suggestion of non-Euclidean rules for computing the distance between objects, Restle also attempted to define similarity as a distance. He defined this measure as the symmetric set difference,

"Other things equal, the degree of dissimilarity between two sets depends on the measure of noncommon elements i.e. the symmetric set difference." (1959, p20)

Restle describes his 'elements' as the "universe of stimulus elements, cue, etc." (p208). Unfortunately he does not regard the case of 'other things not being equal', but would appear to have been aware of the important philosophical problem of distinguishing between 'essential' and 'accidental' features of an object. Tversky's (1977) 'contrast' model, which ignores the problem, was the first to provide an index of similarity incorporating both

common and uncommon elements. Restle notes that Bush and Mosteller's index cannot be a basis of a distance measure because it is not generally symmetric, although the mean of the two directional n's would be. An index equivalent to this mean was suggested by Galanter (1956). His distance measure is based on the ratio of the measure of the set difference to the measure of the union of the two sets,

$$G_{AB} = \frac{m(A) + m(B) - 2m(A \cap B)}{m(A) + m(B) - m(A \cap B)} \quad 2.44.3$$

Restle argues for a generalisation of this equation which has the "desirable properties of additivity when sets are ordered" (p219),

$$D_{ij} = m[(S_i \cup S_j) \cap (S_i \cap S_j)] \quad 2.44.4$$

$$m \geq 0$$

The 'contrast' model of similarity judgement proposed by Tversky (1977) is a further generalisation of Restle's model with the addition of extra weighting parameters, whose significance is detailed below.

$$D(i,j) = \alpha f(i-j) + \beta f(j-i) - \theta f(i \cap j), \quad 2.44.5$$

$$\alpha, \beta, \theta \geq 0.$$

where f is a non-negative scale and α, β, θ , are parameters which determine the weights associated with the model components.

Tversky also describes another possible matching function, the 'ratio' model. With parameter values of 1 and 0 this is equivalent to Bush and Mosteller's (1951) model, and is evaluated in Tversky and Gati (1982).

$$S(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)}, \quad 2.44.6$$

$$\alpha, \beta \geq 0$$

The contrast model

In his major review paper of 1977, Tversky argued that other than in psychophysical direct magnitude estimation studies, the response processes involved in judgements of similarity are not as 'simple' as assumed within the scaling approach. Rather than thinking in terms of dimensional attributes as the basis of similarity, Tversky proposed that a stimulus be considered as a set of aspects or features, and then defined similarity in terms of common and/or different features. He intended that the 'contrast' model be,

"- used to uncover, analyze, and explain a variety of empirical phenomena such as the role of common and distinctive features, the relations between judgments of similarity and difference, the presence of asymmetric similarities, and the effects of context on judgements of similarity." (1977, p327).

Tversky considers that 'recognition, learning, and judgement presuppose an ability to categorize stimuli and classify situations by similarity'. His perception of the fundamental nature of this notion is indicated in a quotation from Quine,

"There is nothing more basic to thought and language than our sense of similarity; our sorting of things into kinds." (Tversky, 1978, p79)

The 'contrast' model assumes a two stage process, although only the second is developed. In the first stage, some internal representation of an object as a collection of features is formed by a prior process of extraction and compilation,

"When faced with a particular task (e.g., identification or similarity assessment) we extract and compile from our data base a limited list of relevant features on the basis of which we perform the required task.

- the term **feature** usually denotes the value of a binary variable (e.g., voiced vs. voiceless consonants) or the value of a nominal variable (e.g., eye color)." (1977, p329-330)

Tversky's implicit perceptual process is effectively a 'bottom-up' processing of the 'complex' into a small number of binary attributes. This contrasts sharply with the 'apperceptive' processes involved in models such as that proposed by William James; where some form of guidance of, or 'feed-forward' to, the perceptual processes by judgemental processes is developed. Tversky only develops and tests the second stage of the contrast model, the possible operations of the three parameters of equation 2.44.5 in matching object features. The theta parameter is concerned with features shared by objects *i* and *j*; alpha with the features of *i* that are not in *j*; and beta with those in *j* that are not in *i*. Tversky hypothesizes that in the assessment of similarity subjects will attend more to common features, whilst in the assessment of dissimilarity subjects will attend more to their distinctive features. Thus the relative size of the theta parameter may be used to model differences between similarity and dissimilarity judgements. In contrast, the 'diagnosticity' or salience of features is reflected in the relative size of the alpha and beta parameters. This leads to asymmetric judgements whose direction is determined by the relative salience of the stimuli. The less salient stimulus is seen as more similar to the salient stimulus than vice versa. Tversky has argued that the asymmetric effects explained by the contrast model are equivalent to Rosch's (1975) theory of prototypes, "- the variant is more similar to the prototype than the prototype is to the variant, because the prototype is generally more salient than the variant." (1977, p333). However a set-theoretic formulation requires that the concept of prototypicality be treated in a relative sense, where objects are defined as being more or less prototypical than each other. This interpretation is not accepted by Rosch (1978, p31).

Tversky makes wide assertions as to the power and generality of the 'contrast' model of similarity; even going so far as to claim that it can aid in the comprehension of linguistic metaphor. These are strong claims for what is effectively a three parameter matching function. The processes that might be considered to be of most psychological interest are 'pushed back' into the prior perception of features; as are the true problems of representation. For instance, Tversky's claim that dimensional variables such as loudness might be represented as a sequence of nested sets is not convincing, given that he has already argued against the distance models on the grounds of the unrealistic nature of the judgement that they imply. We will later note Schonemann's (1977) criticism of Tversky's experiments in terms of the testing of 'subjective' dimensions that are defined ad hoc by the experimenter. He shows that given an experimental paradigm in which the experimenter makes an a priori determination of the attributes and their salience, it is difficult to see how functions such as Tversky's could ever be falsified as formal 'theoretical' models.

The critical tests

The most serious challenge to the conceptual basis of the distance models of similarity came in a Psychological Review paper by Beals, Krantz and Tversky (1968). Their fundamental criticism of scaling was that, "- the representation of stimuli by specific types of distance geometry is rarely questioned. The computational methods in use give a 'best' answer, regardless of whether the underlying model is appropriate." (1968, p127). Their intention was, "from the standpoint of measurement theory", to study the "psychological meaning" of the properties of the scaling models, and the "structure of critical experiments that might test them". The conceptual origin of their ideas were the models of magnitude estimation and probabilistic choice behaviour of Luce et al (1963). Luce had claimed that although the assumption that psychological similarities are distances was strong, it was not necessarily true. He considered the 'mathematical tractability' that followed from coupling similarity with choice theory to

be a greater advantage. From this standpoint, Beals et al proposed three defining measures to be used in their critical experiments,

- a) Decomposability: the distance between points is a function of componentwise contributions
- b) Intradimensional Subtractivity: each componentwise contribution is the absolute value of the scale difference.
- c) Interdimensional Additivity: the distance is a function of the sum of componentwise contributions.

Subsequently experimental investigations of the properties of additivity were made in Tversky and Krantz (1969), and of both additivity and subtractivity in Krantz and Tversky (1975). This last paper will be considered in detail as the same stimuli are later to be made use of in the experimental section. The study was concerned with evaluating the psychological dimensions of similarity between rectangles; a set of 17 rectangles being used. Krantz and Tversky considered two alternatives as possible candidates of psychological dimensions of rectangles, the height and width combination, and the area (height x width) and shape (height/width) combination. Objectively, rectangles are uniquely identifiable in terms of either one of these combinations. They found that neither combination of stimulus dimensions were satisfactory, although the area and shape combination was found to fit the data slightly better. Krantz and Tversky claimed as a consequence of the 'interaction between attributes' that they found, that either more complex 'psychological' combinatorial rules must be specified or that a model satisfying additivity and subtractivity must be adopted. This latter course led on to the development of Tversky's 'contrast' model of similarity.

Schonemann (1977) provided an insightful criticism of these experiments. He developed a proof that a simple nonlinear transformation of the 'objective' dimensions would satisfy the test conditions. Schonemann is also critical of their general approach, "Krantz and Tversky first define the **subjective**

dimensions ad hoc and then check on decomposability" (1977, p164, my emphasis).

Takane's (1981) 'maximum likelihood' scaling procedures, which enable the investigator to test various models of measurement error, have already been referenced. Takane tested the original hypothesis of Krantz and Tversky against that of Schonemann, that the perceived shape difference increases as the area increases. (Further details of the experiments, together with illustrations of the stimuli and hypothesis configurations can be found in Section 4.21 and Appendix A). Takane found that although Schonemann's hypothesis fitted some aspects of his data well, it did not capture the way that, "- the configuration not only diverges along the area dimension, but also is curved in an interesting way; curves connecting the rectangles with the same area levels shape like arcs drawn from a common focal point -" (1981, p25). Takane's 'fan' hypothesis closely fitted the data. Moreover he provides the interesting interpretation that it results from the effect of a 'third' dimension, which involves pairs of stimuli being contrasted with all other stimuli; effectively a global context effect.

Axiom systems

It is clear from the review of the general literature how the contrasting models of similarity are crucially dependent upon the axiom systems on which they have been developed. The mass of theoretical exposition and experimental work that has been built up since the original Beals et al (1968) paper can be shown to be insubstantially grounded as a test of the appropriateness of distance models. Following Coombs' general model for the analysis of data, which was discussed in Chapter 1, Roskam (1979) noted that there are two decisive steps in mapping an empirical set into a formal metric set. The first is to convert the elements of the empirical set into data; by assessing certain relations. The second is to propose a representative model and find the best fitting representation according to the model. We have argued already that the first stage, deciding what

constitutes 'data', is psychologically primary but that for many of the theorists in this domain the second stage, of data description, has been the principle concern.

Bush and Mosteller's (1951) choice of an elemental set-theoretic representation is not controversial. The intention was to evaluate operationally their hypothetical similarity index, and not to coordinate it theoretically with any 'objective' stimulus dimensions. However it can be argued that the same mathematical representation in for example Tversky's 'contrast' model is a severe constraint upon the potential cognitive representations that might be modelled. We have shown that the later set-theoretic formulations have traded mathematical completeness for the 'conceptual difficulties' anticipated by Bush and Mosteller. Furthermore, their criticisms of the scaling approach can be shown to be in error in terms of fundamental measurement theory. In discussing the axiomatization of representation systems, Roskam writes "a specific representation is forced upon the data by axiom - a representation cannot be falsified." (p659).

A fundamental problem of measurement is to decide how much incompatibility can be tolerated between data and model before the model, or the data, is rejected. It is for this reason that the scaling methodologies are all based on loss functions such as Kruskal's 'stress', which serve to indicate the representation that best 'fits' the data. The most that can be achieved in terms of **refuting** a particular model is that, "relations that follow from the model are difficult to believe or do not agree with some data or with established theories" (Roskam, 1979, p659). The 'critical experiments' that Beals et al put forward have been designed within a particular representational model that they have proposed axiomatically. It is inappropriate to seek to evaluate one representation model with proofs and tests based upon a quite different set of axioms. The similarity index is a measure of sameness in the 'eyes of the researcher' which may or may not reflect the relation which exists in the 'mind of the subject'.

"The analysis of similarity indices provides another view, - of the data, which must subsequently be interpreted in terms of the structure of reality." (Roskam, 1979, p208).

In psychological terms, this means that if Beals et al's 'component decomposability' into what are effectively unordered sets of the objects' elemental parts is not acceptable, then their critical tests do not have to be accepted.

Other tests

The earlier paper of Beals et al (1968) questioned the underlying assumptions of the scaling or distance models of similarity, and proposed critical tests of their validity. Tversky's major review paper of 1977 introduced the 'contrast' model of similarity, and a series of experiments intended to illustrate its generality. Although we have already questioned whether simple matching functions may provide useful psychological explanations, some of Tversky criticisms are nevertheless well founded. For instance, he pointed out that most 'distance models' of similarity had generally not addressed the problem of the distinction between similarity and dissimilarity judgements in a thoughtful fashion. Similarity and dissimilarity are often referred to as the simple complement of each other, and most of the MDS programs permit this transformation to be made on the raw data. Coombs' view that "(We) may use similarity and dissimilarity at will" (1964, p433, Coombs' emphasis) exemplifies this cavalier approach. Tversky's experiments demonstrated that similarity and difference judgements are not perfectly negatively correlated. This fact had been commented on by Dearborn (1910) and in other earlier studies concerned with the perceived similarity of handwriting styles. However it appeared to have been ignored by the later literature. The feature matching model can account for this effect, in that common features can be given heavier weight relative to distinctive features in similarity judgements than in difference judgements.

We have already noted that Tversky provided evidence of asymmetry in similarity judgements, and put forward a 'focussing' hypothesis, involving prototypes, in which the direction of the asymmetry is predicted from the relative salience of the stimuli. He also deals with the problem of 'context' in making similarity judgements. Tversky writes that similarity depends on context and frame of reference, and that these are taken to "correspond to changes in the measure of the feature space". The example is given of how when asked to assess the 'political similarity' between countries,

"the subject presumably attends to the political aspects of the countries and ignores, or assigns a weight of zero, to all other features" (1977, p340).

Thus the 'contrast' model accounts for context effects on similarity judgements by assuming that the measure or weight given to the various features changes in different contexts.

Besides the effect of such explicit or implicit instructions, the similarity of objects is also taken to be influenced by the 'effective context', i.e. the set of objects under consideration. It is in describing this later effect that Tversky comes closest to committing himself to the specification of any of the pre-judgement (i.e. pre-matching) processes. In discussing the relation of similarity to grouping or clustering activities, he proposes the 'diagnosticity' hypothesis,

"The diagnostic factors refer to the classificatory significance of features, that is, the importance or prevalence of the classifications that are based on these features." (1977, p342).

The example is cited of the feature 'real' having no diagnostic value in the set of actual animals, but "having considerable diagnostic value if the object set is extended to include legendary animals -" (p342). Various

experiments are detailed showing the effects on similarity judgements of manipulating the set of objects; effects which Tversky claims cannot be explained by the distance models.

Krumhansl's (1978) 'distance-density' model, which was developed in response to Tversky's criticism, has already been described. The principle contention of her paper is that all the results obtained by Tversky are equally consistent with her model, and that they therefore cannot be used to invalidate the application of geometric models on experimental grounds. Whilst noting the "extreme flexibility" of the feature matching models, Krumhansl notes that it is necessary to make additional assumptions about how Tversky's 'context' effects actually affect the assigned weights,

"The model itself does not specify what factors influence how the weights are assigned." (1978, p455).

Krumhansl also makes some fundamental criticisms of what are effectively the representational assumptions of the set-theoretic models. She references the work of Lakoff and Rosch, and points out that theories of semantics based on category membership are unable to account for many effects in natural language. Category membership in terms of 'family resemblance', a network of overlapping attributes, may provide a more adequate description than any feature or set of features.

Tversky had argued that the dimensional assumptions might be inappropriate for semantic stimuli which vary in terms of discrete qualitative features. However a difficulty associated with feature-based models of similarity is that it has been found that a given property may be adjudged to be more 'central' or important to the meaning or appearance of one object than another. Krumhansl also points out the known indeterminacy in semantic description that dates back to the work of Smith, Shoben and Rips (1974). She points out the possible necessity to distinguish between 'defining' and 'characteristic' features of an object involved in a similarity

judgement,

"It may happen that a defining feature of an object, while necessary, may be less salient than a characteristic feature of the object."

(1978, p460).

As Krumhansl notes, theories of semantics based on category membership are unable to account for the 'fuzzy' character of category boundaries, and that the equivalent set-theoretic similarity models will therefore also be inadequate. Rosch (1978, p31) also makes a similar comment, although less pointedly, in introducing a chapter by Tversky describing the contrast model.

In a later paper, Tversky and Gati (1982) report on a series of 4×4 factorial experiments. These involved judgements of binary feature manipulations of plant shapes and schematic faces. The data from these studies is shown to fail the dimensional tests of 'triangle' and 'corner' inequality; axiomatic statements concerning permissible inter-stimulus distances within a dimensional space. They also informally describe an experiment involving six 'highly similar' ellipses and one circle. They found that the ellipses were judged more similar to the circle, than the circle to the ellipses. It is claimed that these results are in contradiction to the predictions of Krumhansl's (1978) distance-density model; although it is not clear that this is so. Tversky and Gati also criticize Krumhansl because she inferred variations in spatial density from MDS analysis rather than experimental manipulation. However it could equally well be contended that their simple manipulations involve a limited perspective as to what constitutes psychological 'space'; and certainly do not constitute a critical test of the distance-density model. It is interesting to note that in this most recent publication, the 'contrast' model is reformulated, and explicitly described, as a **dissimilarity** function. Within a set-theoretic formalism, this is of course merely a nominal change. However it would seem to further indicate the underlying analytic emphasis of the model.

Metaphor and similarity

The treatment accorded to metaphor is a fundamental source of theoretical disagreement between the model to be developed in this thesis and the 'feature-comparison' models of similarity. Any truly general theory of similarity must at least implicitly be capable of adequately dealing with metaphoric relations. We noted earlier in this section, that the proponents of some feature models have explicitly proposed that such models are capable of being extended to provide an explanation of the comprehension of metaphor (e.g. Tversky, 1977). It is contended that such an extension is only possible if metaphor is reduced to a form of 'condensed-simile', that this is invalid, and that there is a **qualitative** difference in the comprehension processes involved between metaphor and simile.

The essential distinction between the various treatments of metaphor (and similarity) found in the literature follows directly from the different possible perceptions of the description and explanation of physical reality. The 'relativist' view is that the objective world is **constructed** on the basis of the constraining influences of human knowledge and language, and is thus not directly accessible. The central, and opposing view, of 'logical positivism' takes reality to be **literally** describable, and assumes that the medium of language can perform this precisely and unambiguously. Orton¹ uses this distinction between 'constructivist' and 'non-constructivist' perceptions of language to relate alternative approaches to metaphor, "metaphor as an essential characteristic of the creativity of language; and metaphor as deviant and parasitic upon normal usage" (1979, p2).

Current theories of similarity implicitly, or explicitly in the case of Tversky (1977), adopt the 'comparison' view of metaphor. This may be considered as a special case of the alternative 'substitution' view of metaphor, and means that metaphor is considered as no more than condensed-simile. Referencing Tversky's argument, McCabe declares the concern with the contrast between metaphor and simile to be futile (1980,

p11). However her main finding is difficult to accommodate with the treatment of metaphor accorded by the feature models of similarity, "Contrary to much prior research on metaphor, context is paramount in determining the quality of a metaphor - the quality of a metaphor is essentially unrelated to the similarity of its tenor and vehicle." (1980, p128). We would argue for a conceptual distinction between metaphor and simile and a reanalysis of McCabe's data by the author lends support to this supposition (details are contained in Appendix F). A comparison was made of the correlations between her subjects' evaluations of the similarity between a metaphor's tenor and vehicle concepts, and judgements of its quality in and out of context. For these tasks, there appears to be a distinct difference between the operations that her subjects made on the marked and unmarked metaphors.

The model of similarity developed in this thesis corresponds in several important aspects to Black's (1962) interactionist treatment of metaphorical understanding. It is grounded in the same relativist conception of knowledge, with the same emphasis on the notion of 'dynamic systems' of ideas, rather than them being considered only as lists of 'static' features. Black's **functional** analysis of how strong metaphors work is essentially a development of I.A. Richards' (1939) 'interaction' view. Black describes the representational aspects of strong metaphor as cognitive devices for 'showing how things are', and describes the **interpretive** response required of the reader in order to realise what lies behind the words used. The fundamental correspondence that can be made with the proposed functional model of relational judgement, concerns the processes involved in the estimation of the worth or 'goodness' of a metaphor. Such an evaluation may be viewed as being a function of the difficulty involved in **creating and running** the components of the metaphor. The sudden onset of comprehension is taken to be indicative of an abstractive, 'holistic' process rather than a consciously analytic feature match. The complex semantic nature of the elements would appear to rule out the fast but trivial perceptual models akin to template matching.

Discussion of the low level feature matching properties of metaphors and their comprehension misses the vital spirit of metaphor, both in its creation and understanding. Matching the components of an analogy is equivalent to 'feature' matching, matching the 'components' of a metaphor is equivalent to merging dissonant 'frames' of knowledge. It is not possible to understand or evaluate a metaphor by analytic decomposition. However just such a 'breakdown analysis' of analogies is their most useful function. In putting forward his 'feature' theory of similarity estimation, Tversky likens analogical and metaphoric processes. It was noted earlier, that the essential simplicity of the treatment accorded to metaphor perhaps indicates the superficiality or empty structure of the feature analytic theories. All the real problems of interpretation and evaluative judgement concerning metaphor are avoided, "The nature of this process is left to be explained" (Tversky, 1977, p349). It seems quite inappropriate, at least to the present author, to consider the comprehension of powerful metaphors simply in terms of the matching of features in an asymmetric similarity function.

2.45 The concept formation literature

It was noted earlier that the 'universal' nature of similarity has caused it to be differentially defined and used in many sub-domains of psychology. The major source for this review has been papers, explicitly directed at some aspect of 'similarity', from 'The Psychological Review' and 'Psychometrika'. However within the 'Journal of Experimental Psychology' literature concerned with perception and concept formation, there exists an equivalent concern for such issues as 'features', 'attributes', 'dimensions', etc. With the significant exception of Tversky's set-theoretic work, there is little cross-referencing between these two domains. Within this sub-literature there exists a constant debate concerning the usefulness of the feature-dimensions distinction. Gardner was one of the first to note that although in the abstract it is easy to maintain the distinction between dimensions and features,

"- the nature of stimulus generation does not completely determine the nature of the processing: dimensionally defined stimuli can be processed as features, and feature-defined stimuli can, under some circumstances at least, be processed as dimensions." (1978, p105)

Interestingly enough Brooks (1978), in the same volume, also proposes an analytic-synthetic processing distinction. He takes analytic concept identification to mean a "- process whose direct effect is to separate aspects of the stimulus and evaluate their ability to predict category membership. The product of such a process is what is commonly referred to as a **rule** -". Non-analytic concept identification involves the category membership of an item being "- inferred from its overall similarity to a known individual or low-level cluster of individuals, where similarity is judged on the basis of aspects or configurations of the stimulus that are **not** weighted for their criteriality for the particular concept being considered." (1978, p180). Brooks distinguishes between his use of the term 'non-analytic' from 'configural'. Configural judgements are taken to use the same type of information as analytic judgements, whilst non-analytic judgements are effectively 'similarity to special cases' and use different forms of information. This theory resembles Torgerson's (1965, 1983) 'Ideal Type' model; of similarity to prototypes.

Our main argument against this whole sub-literature concerns their use of the term 'dimension'. On examination of the experimental material which is intended to distinguish in some way between the processing of features and dimensions it is seldom found to use other than 'bipolar' dimensions. Some researchers (e.g. Brooks) explicitly claim that their results might be extrapolated to a 'continuum upon the dimensions'. However most apparently fail to realise that their implicit redefinition of 'dimension', by choice of experimental material and design, renders unrewarding any critical comparison. Rosch (1978) claims that "features may be defined as pseudodimensions", and instances 'automatic' and 'standard' as levels on the pseudodimension 'transmission'. She continues,

"When the perceiver determines whether he will treat an attribute as a dimension or as a feature, however, he is constrained in his processing of the stimulus by the information properties of those modes" (p75).

However by making it difficult and unrewarding for subjects to process such 'dimensional' stimuli in anything but an analytic fashion, these experiments would appear to minimise any possible processing distinctions. In the next chapter, it is intended to discuss some of the representational aspects of similarity judgements. It will be argued that 'analytic' judgements may be considered to be concerned with the 'features' of objects, whilst 'synthetic' judgements are concerned with relational or dimensional aspects. It would seem obvious that any study in which the subject is aware that only binary attributes are involved, whatever else it is doing, cannot claim to be fully tapping the potential for relational thought. Also, although Brook's subjects performed above the level of chance, he reported that their response to the experiment was one of "- giggles and irritation together with a protest that they didn't know what they were doing." (1978, p172). Judgements of similarity in 'natural' circumstances do not produce this response.

2.5 Summary

It has been noted that the concept of similarity occurs in various roles and in many fields of the psychological literature, although both the previous review of the psychological literature and this critique concentrate on its role as a measure of the individual's cognitive structure. We have shown that there is some confusion in the use of the term, and that it is generally taken to have only a restricted meaning. In Section 2.2 we outlined the problems consequent on its present more limited meaning in general speech, but the restricted use in the literature must also be partly due to the added 'experimental appeal' of the more easily mathematised analytic perspective. We noted in Chapter 1 the means-end reversal of

experimental methodology being developed around mathematically tractable theories. Evidence for this in the origins of the content models can be found when Restle (1959) writes of his desire to "help in unifying the mathematical approaches to psychological problems", and that "the categorical approach has a firmer logical foundation and the advantage that raw data in psychological experiments are usually categorical". (p219)

Inappropriate mathematisation is a phenomenon that has occurred repeatedly in the history of psychology. In 1929, Boring claimed of Herbart's work that it, "- exhibited the not uncommon case in science, in which inadequate data are treated with elaborate mathematics, the precision of which creates the illusion that the original data are as exact as the method of treatment." (referenced in Flugel, 1933, p16) After detailing some of the problems here, we shall attempt to put forward tentative solutions in the development of a more general model of similarity estimation in the next chapter.

An essential distinction can be drawn between the categorical and the relational representations of knowledge implicit in the various theories of similarity judgement. Shepard et al criticises the 'second order isomorphism' ideas of structural models of internal representation, where "any structure assigned to individual units of information or internal representations themselves generally take the rather restricted and primitive form of a **simple list of discrete attributes**." (1975, p125).

Our contention is that the reductionist approach to similarity estimation, exemplified by the 'contrast' model, is misdirected. Realistically complex similarity judgements are likely to involve qualitatively different psychological processes than those that might be sufficient for 'matching' simple single items. It is argued that the implicit mapping of the internal representation of knowledge of the attribute models provides an inadequate framework for any possible extension to more complex structures or processes.

The 'content' models effectively 'simplify' the judgement task to a problem of **choice and weighting** amongst a series of binary attributes. Their main concern has been the derivation and testing of simple combination rules, and little attention is paid to the **meaning** of psychological similarity; stimulus similarity is accepted a-priori in information theoretic terms. Whilst the mathematical proofs are extremely elegant, the simple representational notions underlying them are relevant only for decompositional forms of analysis. Similarity is taken to be a purely **intensive** relation concerned with the classification of sets of objects in terms of their properties or relations. Despite Tversky's claims, the 'contrast' theory of similarity is unable to deal with the more complex **extensive** relations between concepts, or with the possible complex relational processes involving them, e.g. metaphor comprehension. The criticism is made that the 'content' models are effectively proposing a radical increase in the mathematical complexity of models of similarity judgements, with no commensurate increase in psychological validity.

Hierarchical or additive clustering schemes are repeatedly posited as superior alternatives to continuous dimensional representations in that they may more fully reveal the discrete or categorical nature of data. There are obviously many instances where a hierarchical representation will result in the minimum loss of structural information. However, an instance is detailed in the applied studies of Chapter 5, where even in such a case a dimensional representation, or the difficulty for subjects to conceive of the stimuli dimensionally, may still be psychologically the most interesting analysis. In terms of their formal mathematics, it is clear that both dimensional and hierarchical formalisations can be equally well mapped into a theoretically more general network model. However it is still a matter of judgement which is the best representation for a particular psychological process (Holman, 1972). Rips et al (1973) suggest that it is not the representational aspect of these models that may distinguish them, but the types of processing that they imply. We shall take up these issues in the next chapter. The debate has, in any case, been softened somewhat by the recent development of

procedures that combine the most attractive aspects of both representational systems. Shepard writes, "- the different methods of analysis may be better suited to bringing out different but equally informative aspects of the underlying structure" (Shepard, 1980, p397). This open-minded approach is obviously attractive, but suffers from the inherent danger that the analysis becomes mere data-reduction or a process more akin to 'picture-painting'.

Beals et al (1968) claimed that the geometric models of cognitive space are too simplistic, and that the MDS models may adequately represent only the most simple associationistic models of memory. It is obviously naive to think that the **totality** of complex psychological processes might be reduced to only two or three 'dimensions'. The view that the dimensions (or factors) thrown out by the various data analysis methodologies reflect some fundamentally 'true' psychological dimension, is obviously untenable outside the most basic psychomotor domains. They are essentially a product of the experiment, and are in effect **virtual** phenomena created by the experimental subjects' perception of their task. However within the context of a properly designed experiment, it is possible to usefully manipulate such a small number of global variables, and successfully overcome the 'grain' or level of description problems raised in Chapter 1. Torgerson (1965) first suggested such a 'virtual' dimensional analysis, in studying the effects of experimental instructions upon stimulus material that might have been considered unsuited to analysis by a distance model. He showed that although what he termed 'cognitive similarity', "- do(es) not appear to be inherently spatial or dimensional in nature - it does seem that such structures can always be imbedded in an appropriate space." (1965, p390).

A danger of course is that such a dimensional interpretation might be invented, and that the results would be expected, but fail, to have the properties associated with distances. The practical usefulness of such 'virtual' dimensions is examined in Chapter 5 in the context of an applied study. This involved students taking part in a microprocessor course being required to make dimensional similarity judgements concerning

important conceptual terms; which might in this case have been more easily manipulated in terms of hierarchic groups. Some appropriate 'movement' on a critical dimension may be more significant, in terms of the purpose of the study and interpretation, than the equivalent detailed change in cluster configurations. Both dimensional and hierarchical representations obviously capture elements of the data that might be missed by the other. The theoretically important issue, in terms of our argument becomes one of determining the possibility of particular individuals in an experiment having been required to 'operate' in an unsuitable 'mode' of thought, and whether this is brought out in the subsequent analyses.

The facility of a similarity models to be able to explain intransitive judgements is considered of some importance. Tversky (1977) and Krumhansl (1978) indicate how the set-theoretic and dimensional models respectively might account for local context effects such as prototypicality. The idea of similarity relations having to be transitive is a consequence of thinking of them in a frame-free context. This of course is entirely appropriate for many of the subject domains that have concerned the psychometricians, but is claimed here to be less valid where general semantic knowledge is involved in the judgement task. Intransitive relations would follow from global context effects where one or more of the referents acting as datum might encourage the use of an alternative reference frame, or even a different processing strategy. In the language of schema matching theory, this would be akin to the effects of an element of 'top down' processing being incorporated into what had previously been solely considered as a 'bottom-up' process. The importance of such context effects is that they could easily be misinterpreted as data generation error. It is often noted that similarity is 'not a unitary concept' but, excepting Torgerson (1965), little attempt has been made to explain why different strategies might be employed. Such 'idiosyncratic' results would tend to be lost as noise, or discouraged by experimental conditions. A problem found throughout the literature is that the methodological constraints imposed on the subjects' judgements interact with the fundamental axioms of the different models.

The last general issue that we will take up here, concerns the theoretical 'status' accorded to the data elicitation methodology. We have already noted whilst reviewing them, that the psychometric models and their notation systems tend to precisely define the data but not the **processes** involved in its creation. The necessity for an adequate account of the subject's perceptual task, must be emphasised. For example, it is a difficult but not intractable technical problem to analyse data containing intransitive judgements, given a realistic prior model of its source. Takane's work using the new 'generation' of theory testing algorithms has already been commented upon, and is obviously an important development in the assessment of data creating processes. He writes,

"- if the data are ultimately to be represented by some model -, it is certainly preferable that the initial data conversion process itself is in some sense consistent with the representation model of the data"

(Takane, 1981, p75).

The judgements required of subjects in psychological experiments are essentially rather simple in comparison with real 'life' judgements. They are **made** apparently more complex by the imposed requirements of speed, accuracy, etc. These can be viewed as cueing biases towards the adoption of an overly analytic approach to the task, with task simplifying heuristics on the part of the subject made acceptable. We will later examine further some of the issues concerned with the perceptual and representational aspects of similarity, as it is an issue that must be taken seriously if new models such as Imai's (1977), which result from the view of perception as a **transformational process**, are to be seriously evaluated. The consideration given to the perceptual aspects of the data elicitation task provides a clear indication of the thought being given to the underlying assumptions of any particular model of similarity.

CHAPTER THREE A model of similarity judgements

3.1 Overview

It is not intended to present a model of similarity judgement in an equivalent sense to the feature decomposition systems because, as we have previously argued it would not make sense to formally specify a model in this restrictive sense for what is clearly a complex, emergent process. The problems associated with a too dominant analytic framework for psychological research were introduced in Chapter 1. It was noted that the lure of the analytic 'scientific' method may result in a blinkered approach to theory construction and experimental design. Our intention is to define the proposed model in terms of different forms of operation on particular aspects of object representations. It is possible to describe the nature of the process, and to make formal predictions concerning the consequences of manipulations of important components. In the spirit of the argument thus far, the operation of the model can best be described in relation to the feature analytic models as a comparison between regarding a juggler's clubs whilst in the air, as against being on the ground. In the latter position they may well be easier to specify in terms of detailed attributes such as colour and form, but they have lost their essential relational structure.

In the proposed model, it is suggested that the global properties of the referents are synthetically evaluated in terms of their contextual relations, whilst an analytic 'pattern matching' of local properties is made. The central argument is that similarity judgements must entail both analytic and synthetic components. The consideration of this argument entails an examination of the representation of the objects being considered. The first part of the chapter is therefore taken up with a re-evaluation of the analytic-synthetic distinction, and its interaction with the problem of representation in psychological models. A model is then defined in terms of differential operations on the relational aspects

of the stimuli. Finally a set of testable criteria are developed, and the possibility of their verification is discussed in terms of both experimental and applied studies.

3.2 The conceptual necessity for synthesis

The term 'analysis' means a loosening or dissolving, and by extension, a division of a compound into its structural elements. As against this form of decomposition, 'synthesis' is concerned with the construction of connected wholes from separate elements. Kant first proposed that every analysis depended on a prior synthesis. His claim was that if neither nature nor the analyst has made a prior 'putting together' then there could be no 'taking-apart'. Conceptual analysis was therefore considered to require the pre-discursive 'seeing' of formal structure, within which to make differentiations. Both modern analytic philosophy and cognitive psychology, have tended to either undervalue or even ignore synthesis. However philosophers such as Rosen (1980) are now arguing against the analytic tradition that has dominated twentieth century philosophy. The analytic methodology has been thought to be comparable in clarity and correctness to scientific thought, but Rosen shows that it is just as much a subjective activity as synthesis or intuition. His central claim is that a broader view of rationality must therefore also encompass synthesis and intuition.

Rosen shows that there can be no formal definition of the difference between 'form' and 'structure'; essential components of any model of similarity. This follows from the circularity in the definition of structure as a combination of forms. If a form lacks structure it is invisible to cognition, and therefore the analysis of form must make prior use of the concept 'structure'. To avoid an infinite regress, Rosen argues that we must accept intuition, in this sense of the perception of form, as an **atheoretic cognitive faculty**. The core of any general model of similarity, whether it be defined in terms of analysis or synthesis, must therefore include this

apparently metaphysical faculty; and an infinite level of mathematisation of the more amenable parts of the process cannot alter this. Rosen claims that it is only possible to have a 'notion' of intuition and not a 'concept', or formal theory. A consequence would be that it is therefore only possible to have **notions of similarity**. The attempts in the psychological literature to develop and test rigorous formal models of similarity can only be maintained by restricting the domain of interest to fields where the necessary prior synthesis of structure can be ignored. However, there still remain deep conceptual problems even if the study is restricted to common concrete objects. Since Aristotle's time, it has been known that the problem of 'logical priority' shows that the unity of substance cannot be reduced to the discursive list of its 'element-predicates'. For example, 'man' is not composed of 'animal' plus 'two-footed'. Predication states which properties are 'present' in or belong to, the substance, or which properties are 'absent' or do not belong to it. As Rosen argues, even in apparently the most simple case, it is first necessary to **intuit** the substance in order to then make predicated assertions about it.

Brooks (1978) writes of the 'pervasive contrast' made between analytic and nonanalytic thought,

"When successful, analytic processes are supposed to be precisely focussed and relentlessly powerful; in failure they stand in gross violation of common sense. - Successful nonanalytic processes are often held to be the source of deep wisdom of insight and humane qualities; but in failure they represent stubborn, unenlightened illogicality." (p169).

A rigid distinction between analysis and synthesis is shown by Rosen to have been given up in several areas of the psychological literature. They are regarded only as a difference in stress and purpose, for example the idea of 'analysis by synthesis' is common in models of perception. The essential argument of this thesis follows Rosen in that it is claimed that it is not sufficient to provide **only** an account of the analytic cognitive processes involved in making judgements of similarity.

3.3 The representation problem

Psychologists often fail to consider in detail the logical restrictions imposed by the representation systems upon which their theories are based. Given an unsolved conceptual problem of representation, psychologists may choose to ignore it by restricting the domain of study. It is not our purpose here to offer a solution to the problems of representing knowledge, but only to detail the important aspects of representation formalisms as they might affect relational judgements. For instance it is necessary to define the 'complexity' of stimuli in the development of any general theory of similarity judgement. This can only be developed in the context of an assumed knowledge representation, upon which it is dependent. Within the framework of a set-theoretic conception of memory structure, Zajonc (1955) formally described complexity in terms of the depth of embedding of the successive levels of inclusion of a concept. Deese (1962) provided a contrasting definition within a dimensional model of memory. He regarded complexity as the measure of the amount that an object is 'thought about'; effectively the number of associative relations made to it.

This section is intended to indicate the possible differentiation of the models in the literature, in terms of their treatment of the relational aspects of objects within knowledge systems. The intention being to link this distinction to the equivalent operational distinction between analytic and synthetic cognitive processes. Theories of semantic memory have generally been concerned with providing models of how people search through their structured knowledge to find information. Information has been assumed to be coded in terms of features (Rips, Shoben, and Smith, 1973) or in terms of networks (Collins and Quillian, 1969). At the level of representation these two apparently dissimilar models may be mathematically isomorphic in that net relations can take the form of, amongst others, set relationships (Hollan, 1975). The possibility of distinction is in terms of their operational processes. Whilst it is difficult to envisage how the Rips et al model might operate in other than a strongly analytic mode, the similarity

and proximity relations within networks may more easily be envisaged as involving synthetic cognitive processes.

Critics of cognitive psychology have often claimed that significant problems of knowledge representation are glossed over, and that arguments concerning the differences in psychological theories are conducted at the wrong level e.g. in terms of methodology (Putnam, 1981). It was suggested in Chapter 1 that attempts to describe complex cognitive processes within an overly strict formalism may lead to such a 'wrong level' of description. Lakoff has shown that set theoretic conceptions of object categories are not flexible enough to accord with the way that individuals actually categorise their world,

"On the standard objectivist view, we can understand (and hence define) an object entirely in terms of a **set** of its **inherent** properties.

But, as we have just seen, at least some of the properties that characterize our concept of an object are **interactional**. In addition, the properties do not merely form a **set** but rather a **structured gestalt**, with dimensions that emerge naturally from our experience."

(1980, p122)

Part of our argument against the feature analytic models of similarity judgement is that they lead to the acceptance of just such an over simple model of the attributes of an object.

Qualities and relatives

An important aspect of Aristotle's 'differentiae of Genera' is the distinction made between **relatives**, "- all such things as are said to be just what they are, **of** or **than** other things, or in some other way **in relation to** something else" (Categories, 6a36), and **quality** such as states and conditions. In the modern literature a relation is regarded as a connection, or ordering, of two or more things. It is conventionally thought of as a dyadic propositional function, but may also be used to refer to a propositional function of two or more variables. Quality defines how

objects are constituted, the characteristics that they possess and by which they may be recognised. In dealing with the metaphysical problems of reality, philosophers have conventionally distinguished between primary qualities which exist independently of an observer, and secondary qualities which exist only as content in consciousness; the latter of course being the concern of psychology. Aristotle uses 'quality' as an abstract noun to refer to 'qualities as ascribed to things'. Objects are considered to be described by a qualification word, "By a **quality** I mean that in virtue by which things are said to be qualified somehow" (Categories, 8b25). He distinguishes four kinds of qualities: states and conditions; natural capacity e.g. health and hardness; affective qualities; shape and external form. An important issue that we have already noted, is that Aristotle considered similarity to be a necessary aspect of quality; and used this to distinguish 'quality' from 'substance' and 'relatives' in his theory of categories.

Within this delineation of category systems, Aristotle quite clearly accepted that something might be considered either as a quality or as a relative, depending on the way that it was being considered: "Moreover, if the same thing really is a qualification and a relative there is nothing absurd in its being counted in both genera" (Categories 11a 37). In the modern psychological literature, Miller and Johnson-Laird (1976) also attempt to distinguish between the relations and properties of objects. Whilst noting that relations may be thought of as properties, they comment upon the difficulty of actually defining them at all, "What it is that makes a relation seem natural is even more slippery to define than the naturalness of a property - we must rely on intuition to suggest the restricted notions of 'property' and 'relation' that we require." (p323)

Essential and accidental properties

Aristotle first formalised the 'common sense' notion of the distinction between essential and accidental properties of things. The philosophical debate has been continued in terms of attempts to define the 'internal', or

essential, and the 'external', or accidental, properties of objects. Related ideas also occur in the psychological literature. Following Miller and Johnson-Laird (1976), Osherson and Smith (1978) distinguish between a concepts **core** and its **identification procedure**. The core is taken to be "concerned with aspects which explicate its relation to other concepts, and to thoughts, while the identification procedure specifies the kind of information used to make rapid decisions about membership" (p318). This idea of a context-dependent 'core' which serves most, but not all, memory requirements is also found in other theories of category development.

The holistic tradition and configural properties

Kohler and other Gestaltists argued against the 'bundle theory' view of experienced units being the summation of small elements of experience; for example that a tune is made up of the serial addition of the individual notes. 'Bundle theory' has obviously been a powerful orientating factor for theories of classification and feature recognition, in that it is the essence of ideas of discrimination processes in perceptual theories. In contrast the notion of transposability has been a central theme in the study of gestalt phenomena. Certain forms retain their 'characteristic form' even if all their parts are changed, for example a tune transposed into another key. The similarity between the two forms is seen to involve the set of **relationships** that inhere in the whole rather than in the parts of the tune.

Shepard has noted that a consequence of computer models of abstract information processing, has been the tendency to assign associative connections between largely structureless units of information, "any structure assigned to individual units of information or internal representations themselves generally take the rather restricted and primitive form of a simple list of discrete attributes." (Shepard et al, 1975, p125) In the recent literature, Gardner (1978) has given the most thoughtful consideration to these important representational issues, when he distinguishes between the component and the holistic properties of the

stimulus. The component properties, or attributes, of stimuli are either dimensions or features. Dimensions are defined as variables with mutually exclusive levels, and features as variables that exist or not. This distinction between features and dimensions has already been taken up in reviewing the psychological literature, together with the consequent problems of the use of 'pseudo-dimensions' in experimental practice. The critical point of Gardner's analysis is that the holistic properties of the stimulus may include emergent configural properties, which **coexist** with the component properties of the object.

Procedural knowledge

Discussion of the nature of knowledge in terms of its fundamental attributes may often appear to overemphasise its static aspects. In the debate as to what should constitute the essential object or 'idea' of cognition, it is often argued that its dynamic nature must also be considered. The static-dynamic or content-process distinction can be traced back at least as far as Brentano (1874) within the psychological literature, although it is clear that the debate itself is merely a reincarnation of the old philosophical distinction between 'knowing that' and 'knowing how'. His influential work is seen as the foundation of 'act psychology', as distinct from the then dominant 'content psychology' based on simple associationist conceptions. Brentano conceived of the mind as an active creative agency and contended that the essence of experience are mental 'acts'. True mental activity in Brentano's terms occurs when a person 'sees' a colour; the sensations exist but are themselves not mental. Brentano conceived of the mind contemplating its own activity in that the objects of an 'act' might be another 'act'; effectively the notion of 'process as data'. In a more recent exposition of these ideas, Anderson (1976) regards 'declarative and 'procedural' knowledge as parallel encodings, with distinct cognitive operations on the different encodings.

The necessity for coexisting representations

The point that this brief analysis has been intended to emphasise, is that although it is obviously useful for theorists to set up distinguishing categories within their representation systems, their boundaries cannot be precisely delineated. Furthermore it appears important to the 'operation' of the models that the category distinctions between the representational entities be not too strict; i.e. that there be a plurality of encoding within memory. This point appears to be more fundamental than the obvious fact that a less precisely defined model may be more easily 'tuned' to fit a particular set of data. The most obvious argument is that a multiplicity of possible cognitive operations **requires** theories to invoke 'parallel' representations. However it should be noted that recent discussions of the representation problem within the AI discipline have explicitly tended to blur the distinction between declarative and procedural knowledge. Accepting a duality of encoding, the critical problem for the knowledge representation is considered to be how to make it possible for "shifting between levels of knowledge to occur smoothly" (Winograd, 1975, p209).

Fodor (1976) has discussed in some detail the constraints that possible internal representations may have on particular cognitive processes, and on the general nature of psychological investigations.

"- information must be represented somehow, and some forms of representation may be better than others; better adapted, i.e., to whatever task the organism is engaged in. - The key appears to lie in flexibility. Human beings apparently have access to a variety of modes of representation, and can exert a relational control over the kinds of representation they employ." (1976, p194)

The important point that Fodor makes is that the deployment of computational resources, e.g. the mode of representation of information, and its consequent forms of processing, may itself be viewed as a computational

problem which "human beings are, apparently, pretty well equipped to solve" (p194). The consequences for psychological experiments are obvious, they are likely to reveal only "the subjects capacity to figure out the experimenter's goals -" in terms of "the local special-purpose strategies that subjects devise" (p194). It is the intention of this thesis to explore this interaction between possible representations and processes within the context of judgements of similarity.

In terms of the argument that we are attempting to present it appears most useful to consider a continuum of possible representations. Although it is tempting to attempt to differentiate between analytic and synthetic cognition purely in terms of process complexity, or the abstract nature of the entities being considered, upon considering the possible effects of different representation schemes, it would appear best to consider these processes as different forms of **operation** on different **aspects** of a representation. This cognitive differentiation might be most aptly described in terms of 'richness'; the analytic operations entail simple matching of a limited number of object attributes whilst the synthetic operations give rise to a more complex comparison of objects and the patterns of their relations.

A relational formalism

It is obvious that mathematically the distinctions made between the elements and aspects of the various representation systems are not clearly distinct. There can be no formal distinction between relations and properties in that all properties may be construed as relational properties, in terms of some 'universal', and equally an abstraction of a relation can be treated as a property. Models of similarity must therefore accept a fundamental indeterminism in the mode of description of cognitive objects. However it is clear that properties and relations may usefully be regarded as more or less valid descriptions of how objects may be regarded within a particular

context. Moreover the properties of the representation system have profound implications for the possible cognitive processes that can be made 'within' a particular representation. Figure 3.1 shows an abstraction of a knowledge net, with all aspects of the cognitive objects represented in relational terms. In this context, qualities may be considered as **local** relations, extensional relations such as context as **global** relations, and intensional relations such as **configural** properties as being intermediate on the 'local-global' continuum. Discussion of the 'complexity' of an object within this framework can now be made in terms of the differential regard paid to the various levels of the object's relations. The major implication of this claim is that the local 'properties' of an object are processed analytically, whilst the global relational entailments are acted upon by synthetic processes. This is in contrast to many present associative models of memory in the literature, where strictly deterministic analytic operations are made throughout the relational networks (e.g. Anderson and Bower, 1973). However it would be in accord with 'spreading activation' models of human memory (e.g. Collins and Loftus, 1975), and also other models of complex cognitive activity such as the 'interactionist' theories of metaphor comprehension (Black, 1962); to which we will later attempt to link this exposition.

3.4 A summary of issues

Before giving a definition of similarity in process terms, and detailing how its important components might be tested, this is an appropriate point to sum up the argument thus far in terms of the problems not adequately dealt with by current similarity models in the literature. The general, or in Gregson's terms 'literary', models have been shown to suffer from a lack of precise formulation (cf Section 2.42). In contrast, the formal models can be seen to be making very precise statements about the least interesting parts of the problem. Figure 3.2 is an attempt to make explicit the difference between the classes of models. It is obvious that both entail some form of

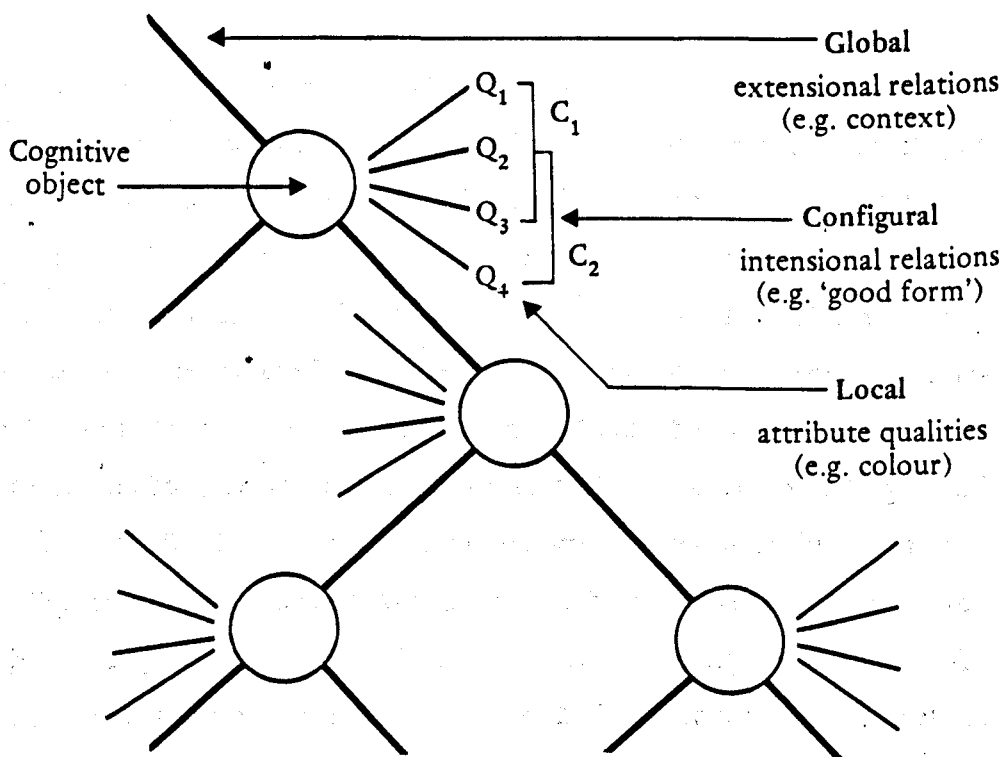
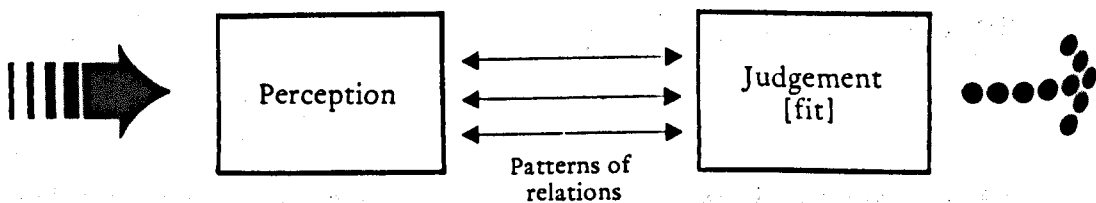


Figure 3.1 A possible network formalism implying local and global aspects

The 'Literary' models



The 'Formal' models

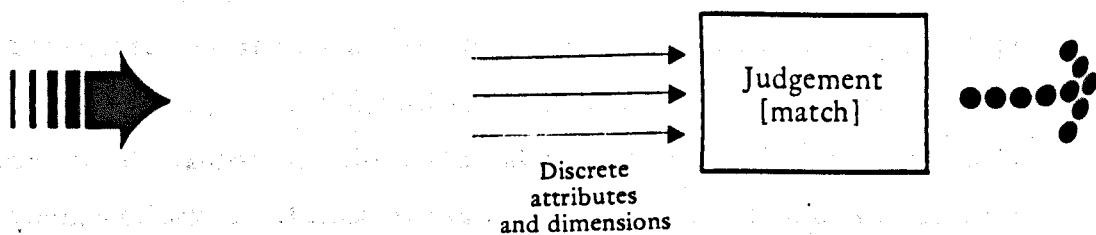


Figure 3.2 A comparison of process implications between the 'literary' and 'formal' models of similarity.

judgement process. In the representational terms that we have outlined, judgement in the literary models entails some estimation of 'fit' between patterns of relations. However for the formal models, judgement becomes a comparatively simple matching operation on discrete attributes or dimensions. It was indicated in the earlier review of the literature that several of the formal models allude to the complex perceptual processes necessary for the existence of the attributes and dimensions that their operation implies. However such processes are always excluded from formal consideration. This of course avoids the necessity to confront any problems of representation. In this respect, it would appear that the 'literary' epithet might be more fittingly applied to the formal models.

There are three important issues that must be adequately dealt with by any general model of similarity:

The true complexity of the stimulus must be accepted in that,

"- a stimulus is not just a vacuous entity; it has properties, and the properties have a great influence on the type of process that can be or is carried out by the humans we use in our experiments"

(Gardner, 1978, p100).

The component properties of the stimulus, whether features or dimensions, are usually all that is considered. However regard must also be paid to the global and configural properties, which coexist with the local component properties. Differences between classes of stimuli that have not been made explicit in a particular model of similarity may make one form of processing more efficient than another. The actual choice of stimulus may predispose the subject to use a particular form of processing, or may limit processing to a particular analytic or synthetic mode. A formal 'mathematical' model is presently only attainable with unacceptably simple models of knowledge representation. For example, the restrictions on the set theoretic representations means that they may only deal with intensional and not

extensional relations. They therefore cannot deal adequately with the development of prototypes, or the relative importance of different aspects of the stimulus; aspects of cognition which have been shown by Rosch (1978) and others to be of crucial importance (cf Section 2.44).

Task definition must be accepted as an important determiner of cognitive processing. The adoption of different cognitive processes, or the differential efficiency of processes, may be the consequence of apparently unimportant aspects of the task set to the subject. For example, constant reference may be found in the literature to the disparities found between judgements of 'similarity' and 'dissimilarity', although a satisfactory theoretical reason for this is seldom provided. (e.g. Schiffmann et al, 1981)

Ignoring the context in which the judgements are made causes the experimenter to miss the true variety and complexity of human cognition. It is necessary to accept the intrinsic conflict between contextualism and systematic description. Although a phenomenon may apparently vary from context to context in a way that apparently defies fixed rules, it is often still possible to derive a systematic model of the concept. As for example, is accepted with the notions of subjective probability and utility within the field of decision theory.

The sum of our argument then is that models of similarity must incorporate some perceptual element, that judgements of similarity must be seen as a balance between analytic and synthetic processes, and that a misplaced rigour in the attempted formalism will tend to be counterproductive in terms of a model's validity and generality.

3.5 A functional model of similarity and its consequent process effects

"In imperfect similarity, however, the only specific implication is - likeness of nature in the relations." Herbert Spencer (1881, p259)

A statement of the model

Judgements of similarity involve an operation of correspondence between objects of cognition, and are always made within some **context**. The operation of correspondence is made in terms of a **projection** of the object's functions within the the individual's meaning space.

Aspects of the model

- The projections of the objects may be viewed as being more or less **complex**.
 - Simple objects would normally be regarded in terms only of their **qualities**.
 - More complex objects would also be regarded in terms of **relations**.
- The **purpose** for which the correspondence is made, is considered to be a facet of the context.
- The context of the correspondence, and the complexity of the objects, will result in a bias towards different 'processing' strategies.
 - The strategies may be thought of as being either **analytic** or **synthetic** in function.
- A consequence of the possible different operations, is that the context within which the correspondence is made determines the relative salience of the qualities and relations which define the individual's perception of the objects.
 - A **local** context would cause only the qualitative nature of the object to be attended to.
 - A more **general** context would also cause its relational aspects to be considered.

A development of the terms of the model

The use of the term 'projection' by Wittgenstein was outlined in Section 2.3. Similarity judgements are taken to be made of an abstract projection of the entities in terms of their functions. It was noted that this notion of projection rules corresponds to the distance models of similarity, and that 'function' in Wittgenstein's terms is equivalent to Herbert Spencer's phrase 'likeness of nature in the relations', which was detailed in Section 2.4.2. We earlier invoked a comparison with a set of juggler's clubs in motion. Function here would refer to the system of relations operating on the clubs. More formally, and within the framework indicated by Figure 3.1, function is taken to refer to the global extensional relations of the cognitive objects such as context.

In his development of a taxonomy of possible cognitive systems, Fodor (1983) uses the term 'isotropic' to describe such global systems of relations. His argument is based on the philosophy of science due to Quine (1964), which demands that due regard be paid to the interconnectedness of ideas. Quine maintained that it is only the total system of assumptions in which a given statement is an element that can be subjected to experimental test. It is therefore not possible to fix the meaning of isolated statements. Our use of the term 'global' is taken to be equivalent to Fodor's phrase 'Quinean or isotropic' in describing the functional interdependence of systems of relations.

The distinction between analytic and synthetic judgements has a long philosophical history, and the necessity to introduce some aspect of synthesis in the perception of similarity has already been discussed. It is a relatively useful distinction found in various equivalent guises throughout the psychological literature, e.g. as visualiser-verbaliser, holist-serialist, leveler-sharpener, etc. Studies using these terms tend to concentrate on aspects of individual differences. Our concern here though is only the facilitation of more or less appropriate cognitive processes by

aspects of stimulus complexity and general context.

The term 'synthetic' is preferred to 'holistic' as it does not have the same image orientated connotations; the terms not usually being distinguished in the literature. Although it must be accepted that analytic categorisation processes are an essential part of human thought, our argument is that psychological models should not concentrate on the analytic to the exclusion of the equally important synthetic and abstractive thought processes. Models of the analytic cognitive processes have had a more respectable allure, as they are amenable to (apparently) explicit formal models and experimental manipulations. Analytic models are also relatively easy to specify in terms of explicit processes, states etc. However we noted in Chapter 1 such problems as that of the 'hidden executive' in AI production system models.

In discussing possible factors that might affect judgements of similarity Tversky refers to "- the effective **context** (i.e., the set of objects under consideration)" (1978, p81). Context effects for the feature analytic models, have only the restricted meaning of the relative salience effects of attribute features; and would be defined as purely local relations within our model. It cannot therefore refer to contextual cueing of extensive relations which we have argued, and intend to prove, must also be taken into consideration. Although Tversky and Gati (1982) use stimuli such as the names of countries, which appear to be conceptually rich in terms of the context that they might imply, the experimental manipulations and instructions effectively cause subjects to concentrate on matching their simple attribute features. (cf Section 4.22 for a discussion of Tversky's experimental instructions.)

We have already referred to the use by the philosophers of the 'qualities' and 'relatives' of objects in their categorisation schemes; and Gardner's (1978) derivation of the equivalent 'content' and 'configural' aspects of the stimulus. A critical assumption of the proposed model is that complexity is a subjective perception involving the intensional 'qualities' and the

qualitatively different extensional 'relatives' of the object. In discussing the problems of testing experimentally the various probabilistic scaling models, Zinnes comments: "For interesting complex stimuli, subjects [perceptions] seem to change systematically over time." (Zinnes and MacKay, 1982, p47). Our argument would be that the subjects are not 'seeing' some new attribute each time that they are asked to make repeated or different judgements, but extending the system of relations that contribute to the meaning of the object for them.

3.6 The operation of the model

An important aspect of the proposed model is not original. We have already noted that the models proposed by Spencer, James and Kulpe implicitly involved a 'top-down' apperceptive component (cf Section 2.42). What is being attempted here, is an examination of the implications of incorporating such a process in both experimental and 'real-life' similarity judgement tasks. This must involve a consideration of the representational aspects of compared objects, particularly in terms of the conditions and expectations set up by actual experiments and task conditions. The most adequate way of describing the proposed model is as we have just outlined, in terms of process effects and interactions. It is sometimes useful to think in terms of the implicit 'executive' of a particular cognitive model (cf Section 1.4). The decision criterion for the the executive of the feature matching, or 'content', models would simply involve the number of attributes to compare, together with some arbitration rules to cope with partial matches. In the same trivial sense that any analogue process may in theory be replicated by a digital simulation, then so it must in principle be possible, at a certain level of reduction, to describe any comparative judgement purely in terms of attribute matches. However we have already noted that in the case of judgements as complex as those we wish to deal with, that this description would then be at the 'wrong level' (cf Section 1.4). It is argued that in executive terms, the model proposed here must

operate at a qualitatively different level and with regard to more complex memory representations. The unit of representation of the process model is assumed to be at the level of complete 'frames'. For example, in the study detailed in Chapter 5, it makes less sense to think of the pen pictures describing the children's school activities in terms of fragments, than as complete situations.

We have already noted that Brooks (1978) provided one of the few reviews of the effects of the dominance of the analytic tradition on experimental design, and details its consequences. He introduced the notion of 'nonanalytic' thought and develops various category formation experiments which facilitate the adoption of such cognitive processes. The important point is made that a wide range of human cognition cannot be described by strict rule systems entailing completely centralized and integrated control. In many cases a general rule can be overridden by information particular to instances. It is interesting to note that whilst developing an analytic-nonanalytic distinction, Brooks effectively distinguishes between forms of similarity judgement,

"- there is no reason why analysis has to be so smart or nonanalysis so dumb. To me, the term **analogy** suggests a nonanalytic mechanism with a similarity device in which there is a control or editing stage to soften the blow of unrestrained physical resemblance" (p199).

To exemplify the possible operation of such a mechanism Brooks instances the case of a chess master not knowing how three possible moves come 'into his mind', but knowing that they must then be analysed in terms of their worth.

It would fit best with the development of terms that we have attempted thus far if 'likeness' were substituted for the phrase 'physical resemblance'.

The term 'resemblance' could then be used to describe the operation of Brooks' 'similarity device'. We might note that here there does not seem to be any important distinction between Brooks' use of 'nonanalytic' and the

term 'synthesis' which has a longer tradition; and also appears to suggest a more positive definition. The more substantial comment could also be made that the analytic-synthetic distinction within similarity judgements should not be arbitrarily assigned as a function of some executive mechanism switching modes, but be viewed as an **intrinsic** function of the representation necessary for particular operations.

Eliminating the executive

We have already outlined a local-global continuum of representation, in respect to which we wish to outline the operation of analytic and synthetic processes. The most simple forms of analytic judgement can be thought of as operating only upon 'local' aspects of the representation. More powerful forms of analytic judgement than that exemplified by Tversky's 'content' model, would also operate on more complex relational aspects, e.g. configural part-whole relationships. The essential operation of synthetic processes can be described in terms of their greater concern for the global aspects of the object; for example the effect of active organizing schemata (Bartlett, 1932) on new items of knowledge.

The crucial distinction made in this model, is that in the analytic case immediate correspondence operations may be made on the 'data' representations, whilst in the synthetic case the 'code' representations must first be evaluated. The synthetic processing of a representation that is not immediately available, but requires some prior evaluation, resembles the workings of Minsky's (1975) frame theory. The necessity for a complex 'unpacking' operation in order to evaluate the relevant relational aspects of an object, is of course the consequence of adopting a serial architecture model of human cognition. The parallel search operation implicit in the 'spreading-activation' models might perform an equivalent function. The argument that Anderson (1983) presents can effectively be simplified to a dichotomy between planned 'serial' operations and automatic 'parallel' ones.

An analytic evaluation concerned with either local or configural aspects might take the form of a simple arithmetic counting function. However a synthetic process would have to be modeled by a more complex proximity operation, functioning in some sense 'within' the representation. As a consequence of these process and representational implications of the different operations, a fundamental distinction can also be made in terms of their relative accessibility to the individual making the judgement. It would be reasonable to expect quite explicit protocols concerning reasoning strategies and the attributes being considered in the consciously immediately available analytic case. However the synthetic operations are only likely to be available for report or introspection in terms of 'feelings' such as those reported by Dearborn (1910) following his 'ink blot' experiments (cf Section 2.42).

3.7 Experimental tests of the model

The essential claim of the proposed model is that the analytic component of judgemental tasks is concerned with the qualitative aspects of the object and the synthetic component with the relative aspects. The most important parameters within the model, and which it is claimed will differentially effect the subjects's cognition, are the 'complexity' of the objects and the 'context' within which the judgements are made. It is possible to formulate a general proposition concerning their interaction of the form,

P1: An adequate description of the discernment of similarity cannot be made in terms of a single simple operation, e.g. the simple matching operation of the 'contrast' model.

Also, more detailed experimental tests of the proposed model are entailed by the following proposition,

P2: As a consequence of content and contextual cues the individual will adopt different processing strategies, and pay differential regard to aspects of the objects being considered.

Two experimental investigations of these effects are detailed in the next chapter, involving manipulations of object complexity and context. A set of three applied studies are also developed in Chapter 5 in order to provide the breadth of evaluation discussed in Chapter 1. The two main analytic criteria used are the analysis of behavioural protocols and a psychometric analysis of the subject's derived cognitive 'space'. Behavioural protocols provide an indication of the information acquisition strategies being employed as a consequence of experimental manipulations. Statistical analysis of the similarities data that is elicited from the subjects provides some indication of the preferred internal representation and cognitive process employed by the subjects. These will be discussed in more detail in the next chapter, in the context of the various forms of data obtained.

4.1 Overview

This chapter details two studies intended to indicate how the feature analytic models of similarity estimation do not adequately describe all possible aspects of the individual's estimation of similarity. The studies were designed to support the argument that a truly general model of such judgements must also encompass a synthetic approach to the task.

The first study was made to indicate the importance of global context for this form of relational judgement. The way in which subjects may create a general contextual interpretation of the stimulus set, as a result of their understanding of the judgement task, is contrasted with the 'local' context effects which are conventionally studied. The latter are only immediately related to the explicit stimulus features and therefore necessitate only an analytic form of judgemental processing. It is argued that a truly general context effect will also involve some change in the perceived importance of the relations of an object, and thus involve some form of synthetic processing. It was noted in the development of the functional model of similarity in Chapter 3, that the feature models cannot account for such perceptual differences, except in an ad hoc fashion by postulating additional feature components in the individual's object representation. The methodological problems consequent on the attempt to control for global context, by effectively setting experimental tasks within contexts that are personally meaningless to subjects, are also discussed in relation to this study.

The second experiment is concerned with similarity judgements concerning two, three and four dimensional data matrices transformed into conventional bar charts or into appropriately graded cartoon faces (Chernoff and Haseeb, 1973). The latter format was considered especially likely to facilitate a

more holistic form of comparative judgements. Use was made of video protocols of the subjects' performance in order to confirm the predicted distinctive approaches being made to the task.

The analysis of both these sets of data is made in terms of their 'fitting', or being adequately described by, the formal assumptions of dimensional models of psychological space. The purpose of these studies is therefore not only to elucidate possible processing and context effects in the similarity judgement tasks, but also to provide a set of theoretical 'benchmarks' for the various multidimensional scaling analyses. Because of the artificial stimulus formats, and the availability of video protocols of performance, it is possible to make use of the analyses of these simple data to 'calibrate', or at least find the most important parameters of, the various analytic techniques. For example, the importance of different levels of 'fit' to prior configurations, and the relative difference of group weightings on the derived dimensions. The more controlled experimental studies may be seen as providing 'external' evidence for the relative facility of the various psychometric techniques used to distinguish between judgements that are 'known' to be the outcome of particular forms of judgemental processing. The same analyses may then be extended to the data elicited in the more complex settings of applied studies.

4.2 Experiment 1 - Context effects

"A satisfactory theory of human cognition can hardly be established by experiments that provide inexperienced subjects with brief opportunities to perform novel and meaningless tasks." Neisser.(1976)

4.21 Rationale

This experiment investigates the effects of global context in similarity estimation. As was noted in Section 2.5, researchers in this field have tended to treat context in only a very limited fashion. It is usually taken to mean simply the presence or absence of a limited number of redundant attribute features. In this study an attempt is made to investigate the effects on similarity judgements of the reference 'frame' that subjects may impose on their judgements in order to make some personal sense of the experimental task. The impetus for this study came from a pilot study using the Krantz and Tversky (1975) rectangles stimuli. Several subjects (who were not psychology students) felt unable to complete the judgement task without first asking what the rectangles 'were'. Dearborn (1910) had reported similar problems with subjects who thought of his ink 'blot' stimuli as 'birds', and wished to make their judgements of them as such. Gregson also refers to the probably under-reported event of subjects being either unable or unwilling to comply with experimental instructions,

"- the studies - have not been consistently successful. - subjects have had to be discarded because they 'did the wrong thing' and followed implicit instructions of their own devising - ." (1975, p66)

The purpose of this experiment was simply to provide groups of subjects with different meaning 'frames', which might be expected to make some personal sense to them, in a controlled fashion. As this stimulus set is ideally suited to the investigation of area and shape biases, frames were chosen

that might be expected to emphasise these effects; subjects were asked to 'think of' the rectangles being used as either skylights or as mirrors.

The Krantz and Tversky study was discussed in Section 2.44 together with the critique by Schonemann (1977), and the reanalysis by Takane (1981). The various constrained solutions for the rectangles stimulus set are shown in Figures 4.1 and 4.2. Krantz and Tversky were concerned that the rectangles should appear 'squat', to overcome a possible area illusion of tall rectangles appearing bigger than shorter ones of equal area. However Schonemann (1977) and then Takane (1981) have shown that the most significant problem with this stimulus set is the effect of perceived shape differences increasing as area increases, as is represented in Figures 4.3 and 4.4 respectively. Full size copies of the stimulus material are contained in Appendix A, but some idea of their relative shape can be gained from Figure 4.5. The stimuli appear more to resemble mirrors and skylights when rotated 90 degrees, and so were used in this orientation for all three groups. The dimensions of the stimuli were created by Krantz and Tversky to fit orthogonal coordinates in log width x log height dimensions. All the studies, including this one, have found unconstrained scaling solutions that approximate to this log x log space. For this reason the later constrained ALSCAL analyses are made in terms of both the actual physical coordinates and also their log transform.

The 'rectangles' stimulus set was designed to test 'additivity' and 'multiplicativity' effects in similarity judgements. Both the original data set, and replications of the experiment, are frequently employed to test the various MDS algorithms in terms of their ability to 'test' various models of judgement. However all the studies with this stimulus set have been partly confounded by the problem of large inter-subject differences. Takane (1981) points out this problem, and circumvents it by requiring a single subject to make paired similarity judgements for several hours a day over the course of a week. The obvious danger with such a long drawn out procedure is that the individual's judgements may be being 'worn down' to a personally meaningless

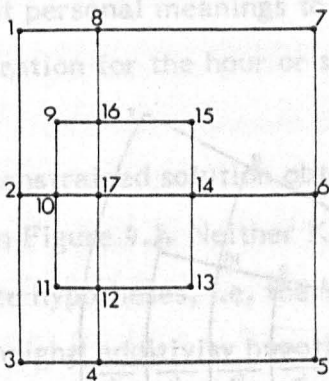


Figure 4.1 Constrained solution under the area (horizontal) and shape (vertical) factorial hypothesis.

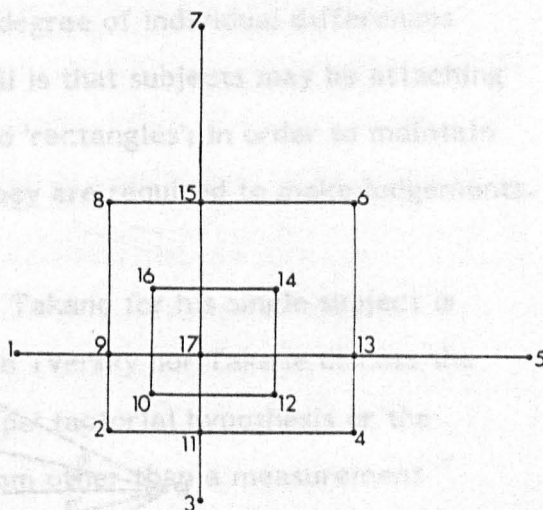


Figure 4.2 Constrained solution under the width (horizontal) and height (vertical) additivity hypothesis.

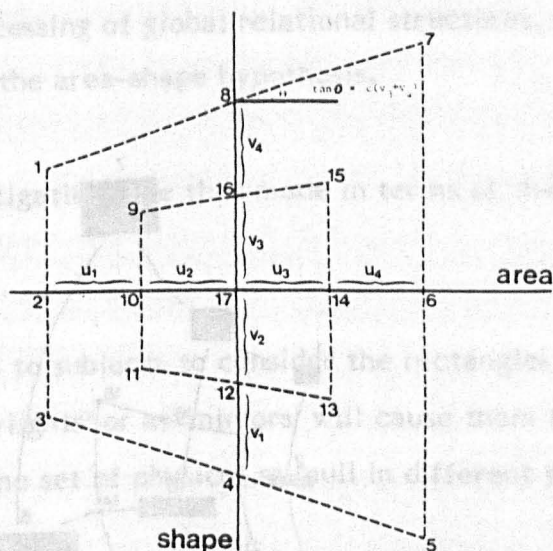


Figure 4.3 Schonemann's hypothesis that the perceived shape differences increases as the area increases

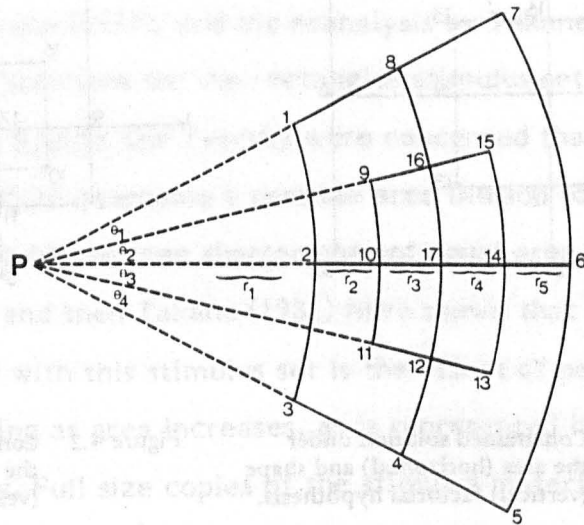


Figure 4.4 Takane's proposed alternative hypothesis to Schonemann's

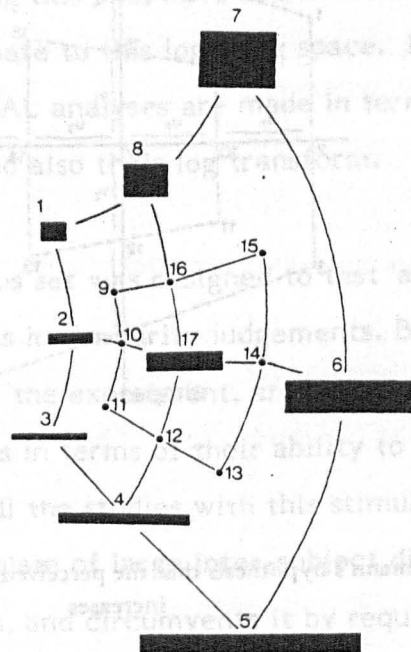


Figure 4.5 Takane's unconstrained two-dimensional solution for his data

consistency. A possible reason for the degree of individual differences found in other studies with these stimuli is that subjects may be attaching different personal meanings to the word 'rectangles'; in order to maintain concentration for the hour or so that they are required to make judgements.

The unconstrained solution obtained by Takane for his single subject is shown in Figure 4.5. Neither Krantz and Tversky nor Takane discuss the alternate hypotheses, i.e. the 'area-shape' factorial hypothesis or the 'width-height' additivity hypothesis, from other than a measurement theoretic viewpoint. However it is clear in the context of our argument, that they could imply quite different cognitive operations on the part of the subjects. Thinking of the rectangles in terms of their component attributes of height and width could be considered equivalent to an analytic cognitive operation, whilst regarding them in terms of their relative ratios, i.e. area and shape, as a synthetic operation. Due to their greater familiarity, the skylight and mirror 'frames' would be expected to facilitate synthetic processing of global relational structures, and consequently fit better the area-shape hypothesis.

The experimental investigations are thus made in terms of the general proposition,

P1: The instructions to subjects to consider the rectangles as 'rectangles', 'skylights' or as 'mirrors' will cause them to consider the same set of physical stimuli in different ways.

and the following experimental hypotheses entailing dimensional salience effects,

H1: Subjects given the instruction to think of the rectangles as 'skylights' will pay relatively greater regard to the area of the rectangles; as would be evidenced by a relatively greater dimensional weight apportioned to the 'area' dimension in a

constrained scaling analysis made under the 'area-shape' hypothesis.

H2: Subjects given the instruction to think of the rectangles as mirrors will pay greatest regard to the shape (i.e. the width/height ratio) of the rectangles; as would be evidenced by a greater dimensional weight apportioned to the 'shape' dimension in a constrained scaling analysis made under the 'area-shape' hypothesis.

H3: Subjects thinking of the rectangles as 'rectangles' would be more likely to make analytic judgements concerning the separate aspects of the stimuli, and less likely to think of the stimuli synthetically in terms of dimensions. This would be reflected in relatively higher 'stress' levels, and lower RSQ values, for the 'rectangles' group in constrained scaling analyses.

4.22 Method

Stimulus material This consisted of 17 rectangles made from thin black card mounted onto small plain white index cards, each completed card being then covered in transparent protective film. A random four digit number was printed on the back of each card for identification purposes. The relative proportions of the sets of 17 stimulus cards were as detailed in Krantz and Tversky (1975). The proportions of nine of the stimulus items can be seen in Figure 4.5, and examples of full size stimulus material can be found in Appendix A.

Design The experiment is of a 'between subjects' design with three groups, the 'rectangles' group, the 'skylight' group, and the 'mirror' group.

Subjects and sessions Five subjects were randomly assigned to each group; this being a convenient number for applying a group correction factor in the intended MAXSCAL-4 analysis (Takane, 1981). The subjects were all second and third year students at the City of London Polytechnic taking a variety of Social Science courses. It was necessary to run the experiment in two separate sessions because of the amount of space required by each subject. The sessions were held consecutively so that subjects might not communicate with each other. Subjects who had to 'sign up' as being available at the time of both sessions were randomly assigned to either session, and within sessions to one of the three groups. They were each paid £2 for taking part in the experiment, which took one hour to complete.

Procedure The three groups were required to make judgements of similarity concerning the stimuli. The 'rectangles' group received a set of instructions which implied that they were to treat the stimuli as 'rectangles'. In the other two groups, subjects were asked to think of the rectangles as representing either mirrors or skylights, the 'mirrors' and the 'skylight' groups respectively. It was emphasized that whilst making comparisons they should think of the 'mirrors' (or 'skylights') as they would normally be used. In the original Krantz and Tversky experiment, a paired dissimilarity measure was used. It is interesting to note that their implicit model of similarity as a form of partial identity is made clear in the experimental instructions used,

"- if the rectangles are almost identical, that is, the dissimilarity between them is very small, mark X -" (1975, p14)

In this experiment, the reference ranking procedure for eliciting similarity judgements was adopted (Young, 1975), and is detailed in the instructions below. This method has been shown by Young to provide accurate and stable similarities data. Even more important for the purposes of this experiment, it was found that subjects much preferred it to the paired similarity task, because it appeared to 'make sense' to them whilst they were carrying it out.

From trial studies it was known that subjects found it difficult to sustain concentration for more than about 50 minutes in this form of task, and that after that order of time the quality of the data obtained becomes so low that it leads to progressively increasing 'stress' levels. Both analysis programs are capable of stable solutions even with large amounts of 'missing data'; and particularly so when the analysis is constrained, or 'fixed', in terms of an initial, i.e. hypothesised, stimulus configuration. Therefore the analyses were made using only six of the seventeen stimuli as referent items, but with all seventeen stimuli present in each case. Each subject received the same item first, chosen from the eleven non-referent items. This 'referent' was intended only for practice, and was not made use of in the subsequent analyses. The following numbers in the referent item list for each subject consisted of a randomized ordering of the **same** six referent items, followed by four others from the subset not used. These were added ^{so} that no subject would feel 'left behind' by others finishing before them. It was emphasized that the subjects should proceed at their own pace, and that they were not expected to finish all the items.

Instructions (The phrases in square brackets represent the different instructions given to the mirror and skylight groups.)

You will find a set of cards in front of you. Each card has a black rectangle drawn on it, and on the back of each card there is a number printed. These numbers have been chosen quite randomly, and are only intended to identify the cards. [You are asked to think of the rectangles as representing mirrors (skylights).] Please spend a few minutes now looking through the cards [, as you do so thinking of them as they would be used as mirrors (skylights)]. I suggest that you make comparisons between the cards^s to see how they vary.

(The subjects were allowed three minutes to look through the cards here, and then the randomized sequence sheets were distributed.)

You have now been given a piece of paper containing a list of numbers, and will have to repeat the following instructions several times:

1. You should find the card which has the first number printed on it and place it on the left of your table, with the rectangle [mirror (skylight)] facing upwards.
2. From the remaining cards, you should choose the rectangle [mirror (skylight)] that you think is most similar to the one that you first put down. [**Please remember to imagine the mirrors (skylights) as they would normally be used whilst you think about them.**]
3. You should now place this card next to the first one with its number facing upwards.

You now have to repeat steps 2 and 3 until all the cards have been arranged from left to right across the table. You should complete the task at your own pace and please remember that there are no right or wrong answers; I am only interested in how you think about it.

(The subjects were then verbally instructed to turn over the reference cards, and to write down all the numbers from left to right in a booklet. They were asked to cross off the first number in the list and to shuffle the cards thoroughly. The subjects were then told to repeat steps 1, 2 and 3 using each number in turn from the list. It was emphasized that they should proceed at their own pace, and that they were not expected to complete the task using all the numbers on the list. Also it was reemphasized that, there were no right or wrong answers.)

4.23 Results

Programs Analyses were made using the general nonmetric scaling program ALSCAL version 4.01 (Takane et al, 1977) and the maximum likelihoods scaling program MAXSCAL-4 (Takane and Carroll, 1981). As with the other scaling programs capable of an 'individual differences' analysis (e.g. INDSCAL), one of the main output plots provided by the ALSCAL program is of individual dimension weightings (e.g. as in Figure 4.6). It is important to note that although stimulus space plots are conceptually in some true dimensional space, the plots of the individual subjects' dimensional weights must be interpreted as the end points of vectors. The angles of the vectors indicate the relative dimensional weighting, and the length of the vectors the average dimensional weighting. The sampling distributions of these parameters is unfortunately unknown. This together with the effects of the complex group normalization and centering operations made on the data matrices have thus far made it impossible to determine a valid significance test for group differences in angular variation.

The RSQ value output by ALSCAL is a group measure indicating the proportion of variance accounted for by the model being tested. This value is regarded as the most appropriate measure of the degree of fit between the hypothesized model and the data, as the various 'stress' measures can be seriously effected by the number of, and the disparity between, data matrices (Takane et al 1977). In the case of the various analyses made here, the RSQ values reported may be interpreted in terms of the proportion of variance in the data accounted for by the particular hypothesis on which the solution is 'fixed' or constrained.

The MAXSCAL-4 program fits the simple unweighted Euclidean distance model by a method of maximum likelihood. It provides several statistical criteria for model evaluation, in particular the AIC statistic; a chi-square goodness of fit to hypotheses representing structural assumptions about the stimulus configuration. The possible analyses using constrained solutions that may be

made with the series of programs are quite unlike those obtained by using 'fixed' configuration analyses with other scaling programs, e.g. INDSCAL and ALSCAL. Equivalence rather than absolute constraints are imposed on the position of the stimuli, e.g. that 3, 4 and 5 should have the same Y-coordinate and that 1, 2 and 3 should have the same X-coordinate. It is therefore possible to constrain the relations of only a subset of stimuli, or even to constrain all but one. (This last form of analysis is made later in one of the applied studies, as it is useful for evaluating hypotheses concerning group differences in perspective of the stimuli subset; particularly where the data is 'noisy' in terms of large individual differences.) Only the SMACOFF programs of De Leeuw and Heiser (1980) also permit such scaling analyses to be made.

The asymmetric chi-square statistic (the AIC statistic) derived from the maximum likelihood analysis of MAXSCAL-4, can be used for the identification of the best fitting model. A large value of this log likelihood statistic indicates a poor fit of the observations to the model. The model is therefore chosen which has the minimum value of AIC. The 'models' in this instance being the hypotheses underlying the constrained solutions e.g. the 'area-shape' hypothesis of Figure 4.1 or the 'width-height' hypothesis of Figure 4.2.

Data Krantz and Tversky produced an individual differences scaling analysis of their data which indicated a wide range of individual differences in terms of dimensional weights. Out of their seventeen subjects, three appeared to be 'using' only the area dimension and three only the shape dimension. The distribution of subject weights for the different dimensions in this experiment is as indicated in Figure 4.6. Of the fifteen subjects only one, who had received the 'skylight' frame, appeared to be making judgements using primarily one dimension; the area dimension in this case. From this figure it is possible to see the difference in terms of assigned dimensional weighting between the 'skylight' and 'mirror' groups. Detailed tables of the ALSCAL analyses are

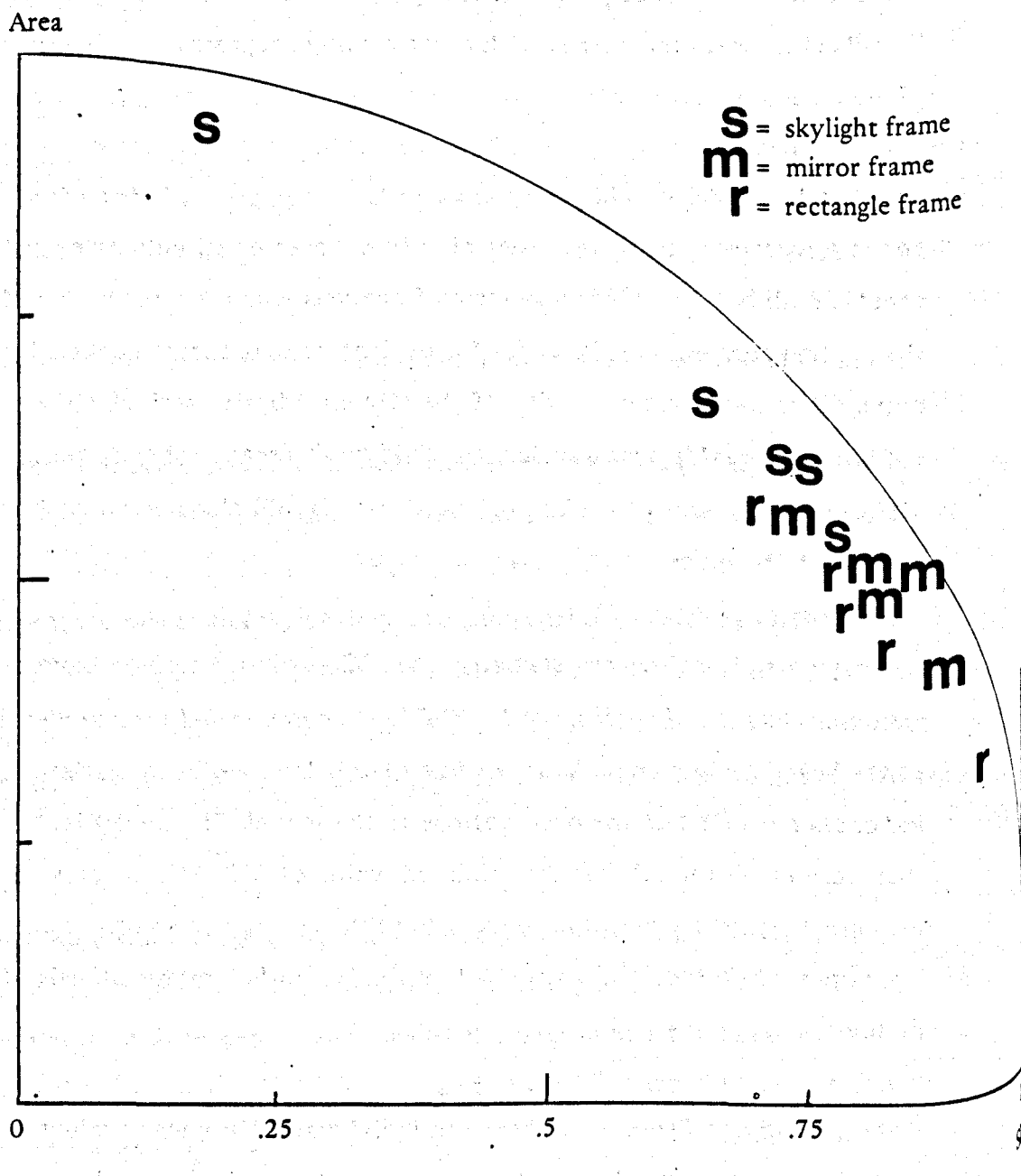


Figure 4.6 Dimension weights from an ALSCAL analysis of the 'rectangles' data

shown in Table 4.2 and that for the MAXSCAL-4 analysis in Table 4.3. A subset of the ALSCAL analyses is contained in Table 4.1 below. This contains the dimension weights, SSTRESS and RSQ values for the three subject groups in terms of the area-shape hypothesis only.

	Dimension weight		SSTRESS	RSQ
	Area	Shape		
Skylight	.69	.61	.17	.89
Mirror	.50	.83	.17	.89
Rectangle	.59	.75	.24	.78

Table 4.1 ALSCAL sub-group analyses - area-shape hypothesis (5 subjects in each group, 17 stimuli, 'fixed' on log x log stimuli configuration)

Each horizontal line in Tables 4.2 and 4.3 indicates a separate analysis, testing the data of a particular group with a particular hypothesis. It was not possible to test either Schonemann's hypothesis, or the distinction between a log and a linear solution with the MAXSCAL-4 program, because of the detailed operation of the 'constrained solution' option. Within each group (i.e. 'skylight', 'mirror', 'rectangle' and combined) an evaluation was made of each possible hypothesis. For the ALSCAL analyses, the results are indicated in terms of SSTRESS, RSQ values and the appropriate dimensional weighting. The results of the MAXSCAL-4 analysis are reported as per the convention of Takane and Carroll, 1981. The log likelihood figures $-2\ln(L)$, are included only for completeness in that model comparisons are made only in terms of the corrected AIC statistic.

As can be seen from Table 4.1, hypotheses H1 and H2 are confirmed in that, as predicted, the average dimensional weighting for the 'skylight' group are highest for the area dimension, and those for the 'mirror' group highest for the shape dimension. It had also been predicted in H3 that

Hypothesis	SSTRESS	RSQ	Dimensional Weight			
			Area—Shape		Width—Height	
<i>Combined groups</i>						
area—shape linear	.270	.751	.53	.73		
area—shape log.	.199	.855	.55	.75		
Schonemann	.230	.819	.60	.70		
width/height linear	.311	.728			.68	.60
width/height log.	.212	.807			.67	.67
<i>Skylight</i>						
area—shape linear	.246	.800	.68	.60		
area—shape log.	.174	.892	.69	.61		
Schonemann	.200	.866	.74	.56		
width/height linear	.286	.771			.78	.49
width/height log	.199	.832			.70	.64
<i>Mirror</i>						
area—shape linear	.252	.775	.48	.80		
area—shape log.	.172	.890	.50	.83		
Schonemann	.208	.853	.59	.75		
width/height linear	.291	.766			.64	.67
width/height log.	.192	.843			.67	.68
<i>Rectangle</i>						
area—shape linear	.308	.675	.45	.79		
area—shape log.	.242	.779	.48	.80		
Schonemann	.276	.734	.56	.73		
width/height linear	.350	.649			.62	.65
width/height log.	.240	.753			.65	.67

Table 4.2 ALSCAL analyses of the 'rectangles' data (17 stimuli, 5 subjects in each sub-group, 2 dimensional analysis)

Group	Hypothesis		AIC statistic
<i>Combined¹ groups</i>	area—shape	—2ln(L)	7761
		AIC	7785
		(d.f.)	(8)
	width—height	— 2ln(L)	7860
		AIC	7883
		(d.f.)	(12)
<i>Skylight</i>	area—shape	— 2ln(L)	2586
		AIC ²	2610
		(d.f.)	(8)
	width—height	— 2ln(L)	2567
		AIC ²	2603
		(d.f.)	(12)
<i>Mirror</i>	area—shape	— 2ln(L)	2402
		AIC ²	2426
		(d.f.)	(8)
	width—height	— 2ln(L)	2497
		AIC ²	2533
		(d.f.)	(12)
<i>Rectangle</i>	area—shape	— 2ln(L)	2642
		AIC ²	2666
		(d.f.)	(8)
	width—height	— 2ln(L)	2727
		AIC ²	2763
		(d.f.)	(12)

¹ These AIC values are not comparable with the sub-group values

² Small sample correction applied; 5 subjects in each sub-group

Table 4.3 MAXSCAL—4.1 analyses of the 'rectangles' data (constrained solutions, 17 stimuli)

the rectangles group would be more likely to perform operations which would effectively be an analysis of the 'separate' aspects of the stimuli. MacCallum's (1981) Monte Carlo evaluations of SSTRESS levels indicate that a value of .23 signifies 'moderate' error. This is exceeded in the case of the rectangles group for both hypotheses, indicating that these data fit a dimensional analysis less well than for the other two groups. Further evidence for the confirmation of H3 is provided by the fact that the RSQ values for the 'rectangles' group are substantially lower than the equivalent values for the other groups. Informal observation of the subjects indicated that subjects in the 'rectangles' group were more prone to make some form of pre-judgemental clustering of the stimulus material. Such a qualitative difference in strategy adopted would explain their poor 'fit' to a dimensional analysis. This observation prompted an analysis of such changes in processing style in the next study; using video recordings of the subjects' behaviour.

The MAXSCAL-4 program does not produce an index of dimensional weighting⁴⁵ such. However this may be adjudged from the relative proportions of the dimensions on the plotted output, in relation to the total 'size' of the output configuration. Figures 4.7 to 4.9 attempt to show this effect for the 'skylight', 'mirror' and 'rectangles' groups respectively. These plots have been superimposed on a log plot of the original stimuli proportions, i.e. upon rescaled versions of the configurations shown in Figures 4.1 and 4.2. A scale factor for the MAXSCAL-4 plots was chosen that would cause the distance between stimuli 3 and 5 of the 'skylight' analysis to 'fit' the underlying representation of the original stimuli. This scale factor was then applied to both dimensions of all three output configurations. The final superimpositions were made by centering the output configurations on the position of stimulus item 17. Also drawn on these plots are ellipses indicating the 95% credible intervals for the positional value of each stimulus point. (The validity of this mapping procedure, and of the consequent interpretation of group dimensional weighting in terms of the

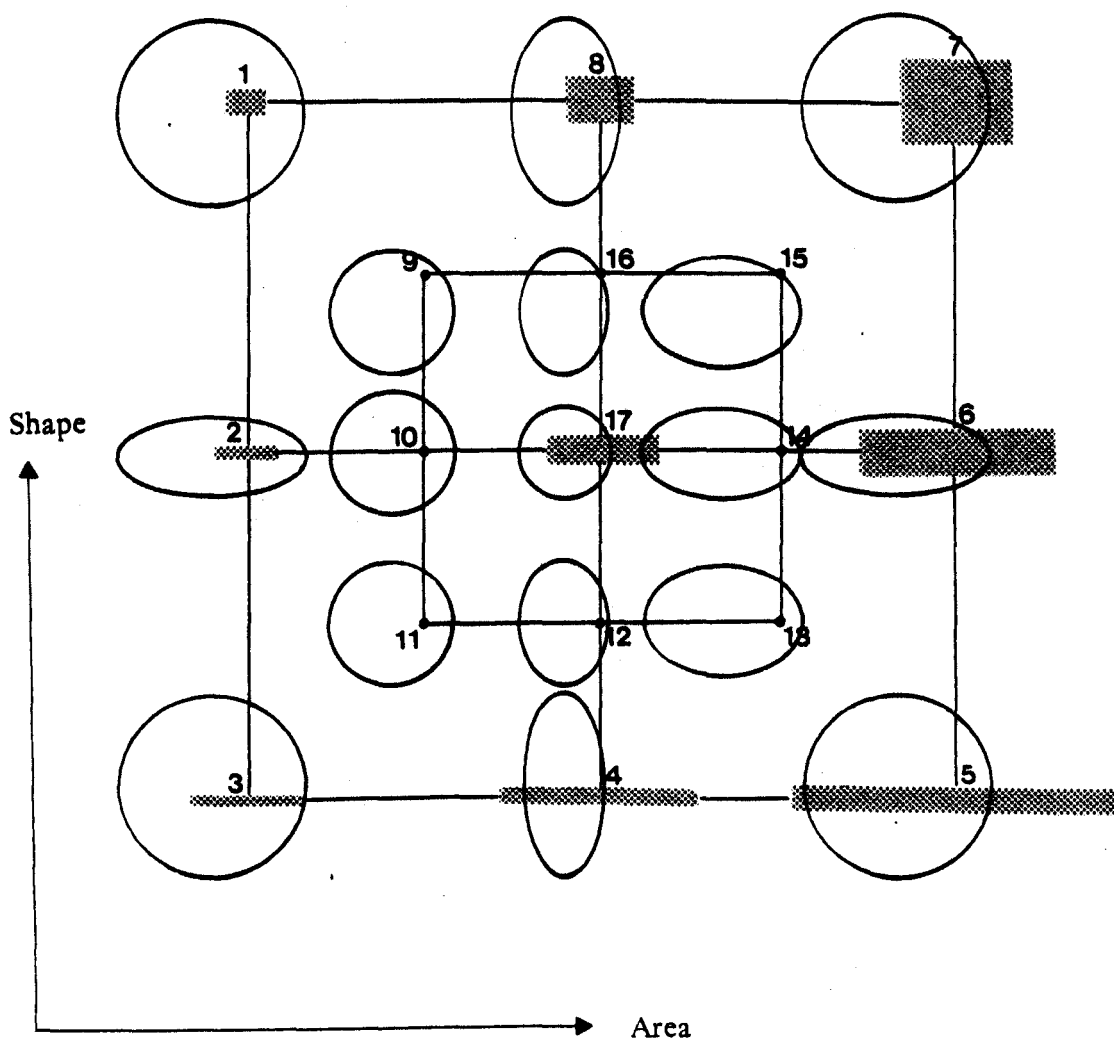


Figure 4.7 MAXSCAL-4 analysis of the 'skylight' group (Ellipses indicate 95% confidence regions)

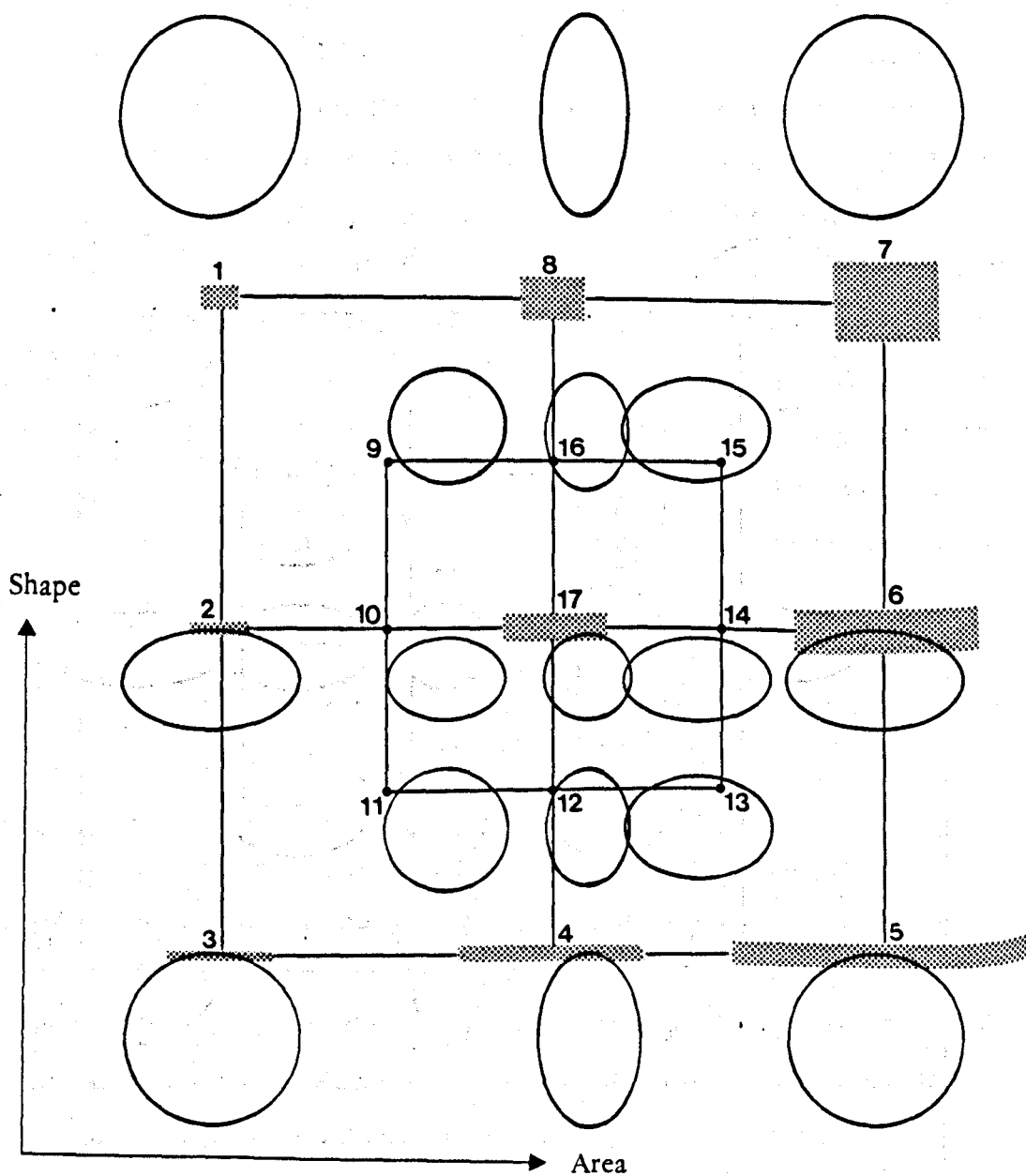


Figure 4.8 MAXSCAL-4 analysis of the 'mirror' group (Ellipses indicate 95% confidence regions)

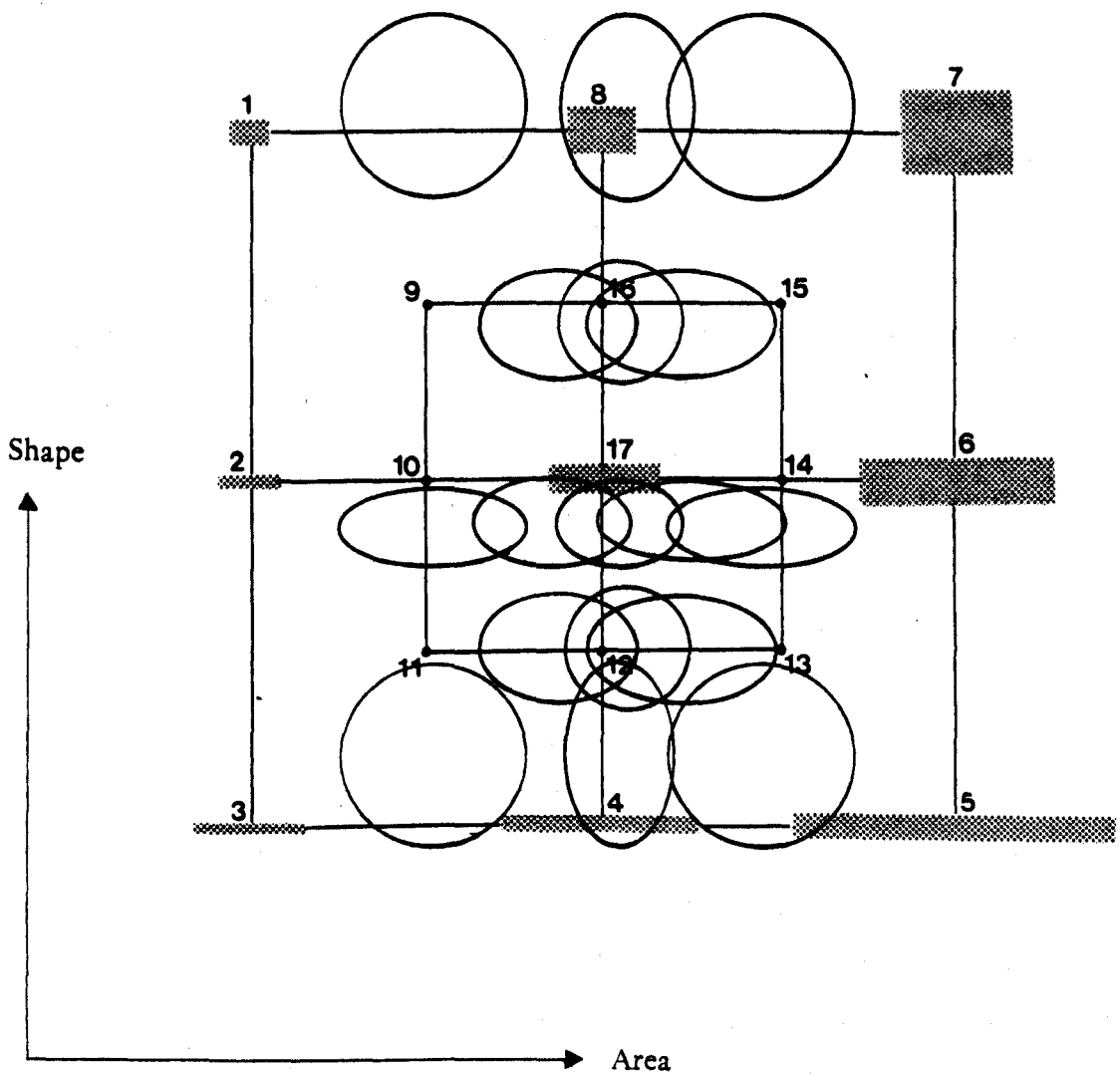


Figure 4.9 MAXSCAL-4 analysis of the 'rectangle' group (Ellipses indicate 95% confidence regions)

relative proportions of the plot dimensions, had been confirmed to the author by Dr Toda in a personal communication, January 1984.)

The analysis of the skylight group indicates an approximately equal weighting between the dimensions of 'area' and 'shape', in that the output plot is balanced in terms of both dimensions of the underlying configuration. For both the mirror and rectangle groups, the 'shape' dimension is given a relatively higher weighting and thus **H1** is confirmed. **H2** is also confirmed by this analysis in that the weighting on the 'shape' dimension is relatively higher for the 'mirror' group than the other two. The ellipses indicating 95% credible intervals start to overlap on the 'area' dimension for the mirror group, and the amount of overlap for the 'rectangle' indicates a large degree of error in terms of dimensional judgement. This provides further confirmation of **H3** in that the 'area' dimension for the 'rectangles' appear to have collapsed.

4.24 Discussion

This experiment was concerned with the evaluation of the general proposition that subjects would consider the same set of physical stimuli in different ways as a consequence of instructions to regard the rectangles in different, and in more or less personally relevant, ways. In that the data was found to prove the three detailed hypotheses, this proposition is taken to have been supported together with the ideas concerning cognitive representation and process in Chapter 3. There the importance for judgements of similarity of the perceived relations of the objects, in addition to their detailed attributes, was argued. The intention of this experiment was to effect general context manipulations that might bring out the claimed differentiation between the analytic operations on the attributes of objects and the synthetic operations concerning their relations. The purely analytic 'feature' models could not account for the perceptual differences found here except in the ad hoc, and hence unsatisfactory, fashion of postulating

additional feature components in the individual's object representation; e.g. mirroriness, skylightness, houseness etc.

It could be argued that the general proposition and these results are intuitively obvious. Nevertheless the implications for studies employing similarity judgements are important. Inter-subject differences, and the intransitivities within an individual's judgements, may not be the consequence of some arbitrary 'noise' parameter, but of the quite logical adoption of some personally relevant global perspective. The important distinctions found between the 'rectangles' group and the other two groups, make it clear that efforts to override the 'world-knowledge' frames that individuals may bring to a task by effectively imposing a sterile framework for the judgements, may be confounded by the individual subject's 'effort after meaning' (Neisser, 1976). These results also have implications for applied experimental studies making use of similarity judgements. In such studies, the referent items are likely to be even more complex than those used in 'pure' experimental studies. An inadequate task framework which does not take account of reasonable interpretative schemes likely to be used by the subject is, therefore, more likely to produce apparently idiosyncratic results. It might be argued that if a pronounced context effect can be induced with such abstract stimuli as these, then the effect would be even greater under 'natural' conditions.

4.3 Experiment 2 - Strategy and complexity effects in similarity judgement

4.3.1 Rationale

The philosophical analysis outlined in Chapter 2, and a reading of the early experimental psychology literature, indicate that making similarity judgements should not be seen as a simple unitary process. It is proposed that the individual may adopt a variety of strategies, ranging from the **analytic** comparison of features of the stimulus, to a **synthetic** consideration of its functional relationships. The actual form of process adopted being dependent upon the individual's perception of the task and the stimuli. The conceptual basis of this study is the 'projection rule' model of similarity due to Wittgenstein. Here similarity is taken to be an abstract projection of the compared entities in terms of their functions. These ideas were discussed in Section 2.3 and are represented in Figure 2.1. There it was shown that particular transformations of letter sequences into a symbolic arrow format might result in perceptually quite different representations. It was argued that the transformations, by different 'grammars' in Wittgenstein's terms, would facilitate particular forms of cognitive processing.

In the notation of Figure 2.1, the 'meanings' in this study are provided by numeric matrices of different dimensionality. The 'grammars' are represented by the transformation of these matrices into the format of either bar-charts or 'Chernoff faces' (Chernoff and Haseeb, 1973); a form of cartoon representation. Examples of both of these formats are shown in Figure 4.10, and further examples are contained in Appendix B. Chernoff applied elaborate transformations to his to-be-encoded data in order to counteract the known biases of human observers for various facial attributes. The dominance of aspects of the eye and central facial features in relation to general facial outline, ear shape and position, etc were important to Chernoff; who was attempting to encode fifteen or more dimensions, but not of significance in our present more limited use of the format. The studies that he details were

primarily concerned with the detection of transient effects in the complex changing patterns of seismic data; often involving twenty or more truly continuous dimensions. Geologically naive subjects showed an ability to detect patterns that would normally have taken several years of training to develop. Chernoff claimed that this was due to their ability to process in parallel the relations between the features of the facial data representation.

Studies using cartoon faces may also be found in the psychological literature, although many involve relatively trivial use of the format. Such studies typically involve only two or three binary 'dimensions', and subjects making paired comparison judgements between stimuli which contain varying numbers of additional redundant features (e.g. Tversky and Gati, 1982). Subjects are even asked to pay close (analytic) attention to the presence or absence of such features e.g. Naveh-Benjamin and Pachella (1982). Such experimental manipulations contrast markedly with those of Chernoff, who was effectively requiring subjects to find discontinuities within an ordered set of data points.

Tversky and Krantz (1969) used three dimensional binary manipulations of faces to test interdimensional additivity effects. As might be expected with such stimuli, the subjects employed clustering strategies and, "support(ed) the hypothesis that a simple combination rule can describe mental processing of independent impressions -" (p127). They attempt to suggest that such combination rules might form the basis for the "interactive or Gestalt effects (that) are important in impression formation" (p127). A more interesting use of the format is found in the study by Takane and Carroll (1981). They used lines of varying curvature to represent the shape of the eyes and mouth on cartoon faces; the stimuli being coded in terms of five different values on each of these two dimensions. In this case the subjects' similarity judgements clearly fitted a dimensional analysis. Takane found two additional statistically significant dimensions which he describes in terms of the skew-symmetry of the mouth and eyes of the stimuli. The

subjects were effectively attributing values of 'happiness' and 'surprise' to the cartoon faces.

In the light of this literature it was predicted that the bar-chart format, with its discrete attributes graphically signalled, would cue and facilitate analytic judgements. The dominant cognitive operation would be expected to be the local processing of the information represented by individual bars in the chart. Conversely the cartoon face format was expected to facilitate synthetic processing. The literature referenced by Chernoff, and his own studies (Chernoff and Haseeb, 1973), suggest that due to the individual's developmental experience, the dominant cognitive operation expected of cartoon faces is in terms of a global perception of the interrelations between aspects of the face. Our argument, and the design of the experiment, is not directly concerned with the relative ergonomic efficiency of the different display formats. However the additional manipulation of stimulus dimensionality is intended to serve as a simple measure of complexity. The correspondence between analytic and synthetic judgement processes and a parallel-serial 'processing' distinction has already been argued in Chapter 3; in particular, the representational implications of 'complexity' in Section 3.5. The literature (e.g. Anderson, 1983) suggests that as the stimulus complexity is increased, it would be expected that the synthetic, or parallel, forms of processing would become relatively more effective.

Observation and interviewing of subjects in Experiment 1, and also in other trial studies, indicated that subjects were adopting different judgement strategies. In this experiment, video recordings were made of the individual subjects' card sorting protocols whilst completing the reference ranking similarity estimation task. If the bar chart group behaved as predicted, then their analysis of the separate aspects of the stimuli should be apparent in terms of their forming and manipulating sub-groups or clusters of items. There is no obvious behavioural correspondent of the hypothesised cognitive operations of the face group; the synthetic interrelation of

aspects of the stimuli. However given the physical constraints of the reference ranking task, it would be expected that subjects making judgements in terms of the cartoon-face format would place the cards directly in some order of dimensional 'distance' from the reference item. It was intended that the behavioural protocols should provide an insight into the judgement processes being employed. Given that group differences in perspective could be made apparent, this would then make it possible to argue more strongly that the form of hypothesised judgement processes can be inferred from the 'fit' of data to the assumptions of the abstract data analysis models.

The transformation of the numeric matrices into bar-chart and cartoon-face format is expected to cause subjects to use different processing strategies for each format, when making judgements of similarity. The dependent measures for the assumption of different processing strategy by the two groups are the classificatory analysis of the behavioural protocols whilst completing the reference ranking task, and the 'fit' of the resultant similarity judgements to a dimensional analysis. The cartoon-face format is expected to facilitate synthetic processing and the bar-chart format analytic processing. It is predicted that,

H1: The bar-chart group will make use of sub-groups and clusters of stimuli cards in the task. The cartoon-face group are expected to order the cards more directly, and accurately, in terms of some dimensional distance; and hence provide data that fits better the original dimensions of the stimuli set when subjected to a scaling analysis.

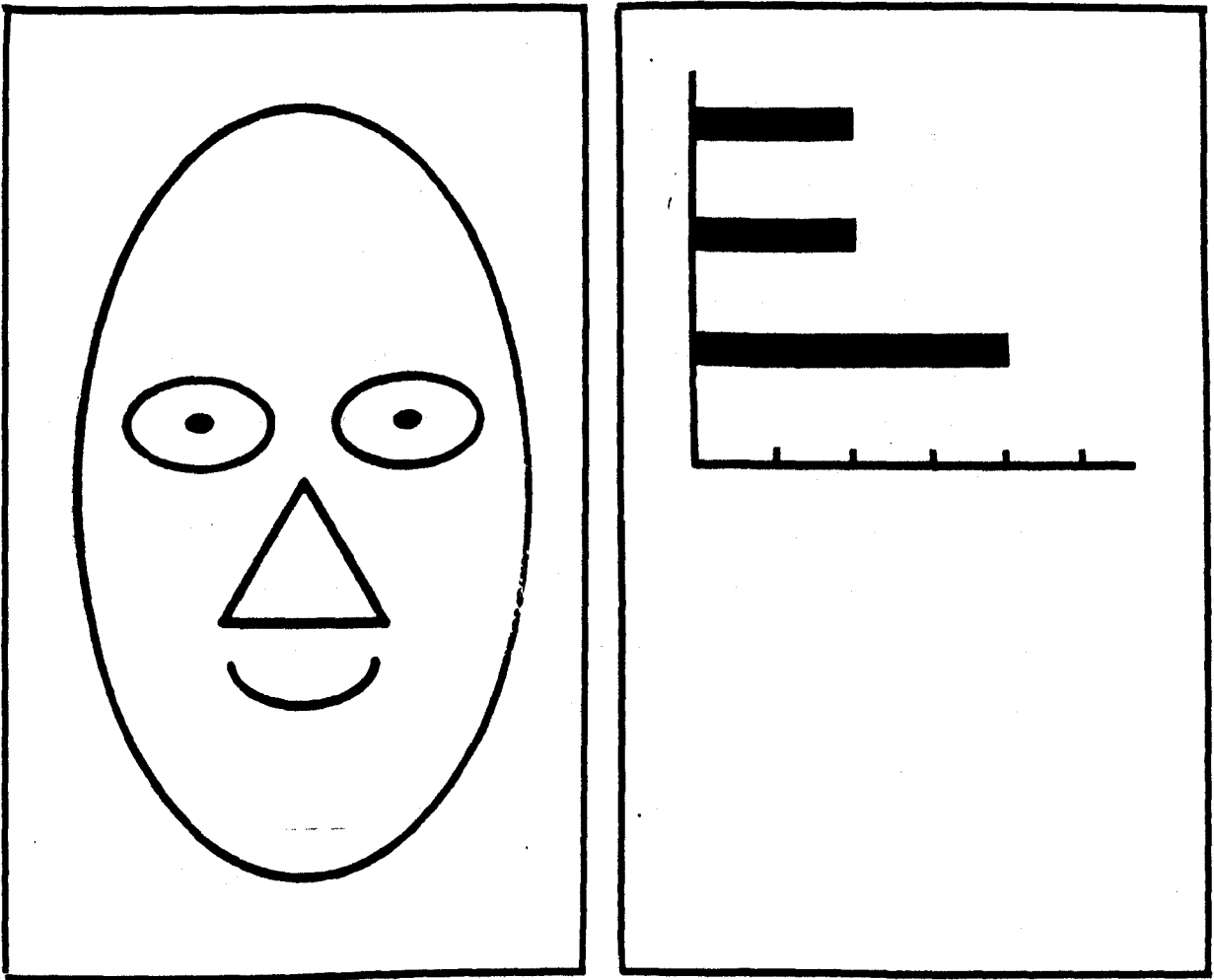
H2: The increase in stimulus complexity, in terms of information dimensionality, will have less effect with the format which facilitates synthetic processing. The cartoon-face format is expected to result in a more accurate mapping of the original data matrix with increasing dimensionality, relative to the bar chart format.

The dependent measures for this hypothesis are that the dimensional weighting following a scaling analysis of the similarities data will be relatively higher for the cartoon face format, as the dimensionality increases, and that the differences in behavioural protocol predicted in H1 will also become more pronounced.

4.32 Method

Stimulus material. Two, three and four dimensional data matrices were encoded into Chernoff face and bar chart format, as shown in Figure 4.10. Details of the source data matrices, and further examples of the encoded formats are contained in Appendix B. Each aspect of both stimulus formats were encoded in terms of five discrete values as Takane's (1981) experiment with similar stimuli had shown that this is a sufficient number for subjects to effectively perceive a variable as a continuous dimension. The cartoon-face and bar chart stimuli were mounted on small white index cards, each card then being covered in transparent protective film. It was of course necessary to have the numbers on the front of the cards in this experiment, and a further index card bearing a random four digit identifying number was therefore attached to each card. This arrangement is made clear in Figure 4.11. Following the 'rules of thumb' of Kruskal and Wish (1978) for the validity of scaling analyses of varying dimensionality, 17 stimuli were employed for the two and three dimensional configurations and 25 for the four dimensional. As in the previous experiment, a 'missing data' design was employed using only five stimuli as the referent items in each case.

Subjects and sessions. The subjects were all second and third year students at the City of London Polytechnic taking a variety of Social Science courses. Subjects were required to 'sign up' as being available for two one hour sessions on consecutive days and were then randomly assigned to the two groups. They were paid £ 2 for each session attended.



Reference stimulus number 15

Figure 4.10 An example of the cartoon face and bar chart formats (The other reference items, and the source data matrices for all the items are contained in Appendix B)

Design A mixed subjects design was employed. One between subjects manipulation was made; the two groups received data in either cartoon face or bar chart format. One within subject manipulation was made; each subject received data at 3, increasing, orders of dimensionality.

Procedure The reference-ranking similarity estimation procedure was used as detailed in Section 4.2. This method had previously been found to provide stable judgements and to make more 'sense' to the subjects than the alternative paired similarity measure. Also it most closely resembles Chernoff's use of an ordered pattern of stimuli, and might be expected to facilitate the synthetic pattern processing that he was able to establish. However its most important aspect in terms of this experiment, is that it provides the opportunity to gain some insight into the subjects' similarity estimation strategies by observing their behaviour as they place, sort and compare the stimulus cards. A video recording of their behaviour was made and a subsequent detailed analysis carried out.

Due to the length of time necessary for the detailed analysis of such video protocols, it was possible only to use five subjects in each group. Experience in trial studies, and the first experiment, indicated that subjects are unable to make reliable judgements of this form for more than about fifty minutes. It was therefore necessary to split the experiment into two sessions of one hour. This included a five minute rest period and five minutes for writing comments in the first session. Each subject was run individually. After the instructions for each subtask had been made clear, the subject was left to complete the reference ranking tasks with only a video camera overlooking the table on which they were to arrange the cards.

The proposed hypotheses are only concerned with group differences between the cartoon-face (henceforth 'face') and bar chart (henceforth 'bar') groups. Therefore 'order' and 'learning' effects were not examined; each subject being given the stimulus sets in order of increasing dimensionality. Within each set the subjects were given one practice referent item, and then

a random ordering of the same five reference items to be analysed. Because of the time constraints imposed by the dual sessions, it was necessary for subjects to make judgements concerning three of the three dimensional stimuli in the first session and a further two in the second session.

Instructions. This set of instructions was arrived at after several trial studies where subjects were later questioned concerning their perception of the task, and how they had approached it. In these trials it was found that attempts to time the students, group trials where they might be made aware of their progress relative to other subjects, and any perception that there had to be a 'correct' answer encouraged analysis of the features of the stimuli and discouraged more holistic comparisons. The use of the 'country' frame for both groups in the experiment was partly intended to aid the subjects in making some consistent 'sense' of the experiment (cf the previous experiment), but also as a constant reminder that all aspects of the stimuli were to be thought of as being equally important. The study by Takane and Carroll (1981) had shown that simple affective interpretations might dominate the judgements of the face group, for example the perceived 'happiness' of the face.

You will find a set of cards in front of you, each with a bar chart [cartoon-face] drawn on it. You are asked to think of each card as representing a country, and that you are being given information about equally important aspects of each country. For example its population, land area etc. Please spend a few minutes looking through the cards now. You will notice that the different aspects of each 'country' vary quite independently of each other.

[The purpose of the random identifying number was explained to the subjects, as being necessary to identify the cards. They were then allowed three minutes to look through the cards.]

Each of the cards has a different number printed at the top. These numbers have been chosen at random, and they are only used to identify each card. You will shortly be asked to take one of the numbered 'countries' and place it on the left of the white card. You should then arrange all the other 'countries' according to how similar you think they are to the first 'country' that you put down. The 'country' that is most similar to the original should be put next to it, the next most similar next to that, and so on until you have arranged all the 'countries' on the white card. You should complete the task at your own pace, and please remember that there are no right or wrong answers to it; I am only interested in how you think about it.

[Because of the constraints imposed by the aspect ratio of the video system, it was necessary for the cards to be ordered in two rows. This was explained verbally to the subjects, and examples of the resulting scene is shown in the photographs of Figure 4.11.]

4.33 Results

A qualitative analysis of the video protocols was made. In order that the hypothesized ranking strategies were not be 'read into' this necessarily subjective task, the categories were identified and 'named' by a research clerk who was unaware of the experimental hypotheses. Cross checking of this analysis was made by an independent judge, and no substantial source of disagreement was found with the original categorisations. There were found to be five clearly differentiable methods of going about the task, which were named 'dealing', 'table sorting', 'scenic', 'grouping' and 'hand shuffling'.

Dealing took place when the subject dealt the cards onto the table in their original sequence, without apparently studying the cards at all. The actual arrangement of the cards in relational sequence did not begin until

all the cards were in view on the table.

Table sorting involved the subjects spreading all the cards out onto the table, then placing individual cards in sequence without any obvious overall strategy. The subjects did not appear to make explicit comparisons between pairs of cards.

Scenic arrangement involved the subject putting the cards down at some 'judged' distances from the reference item; apparently depending on how similar to it they thought each card to be.

Grouping occurred when cards were placed in groups with other cards having apparently similar properties, before being placed in final sequence.

Hand shuffling involved the subjects sorting through the cards in their hands, before laying any card down on the table.

Examples of the 'scenic' and 'groups' strategies are shown in Figures 4.11 to 4.13. Figures 4.14 and 4.15 indicate the different strategies employed by the two groups as they approached each sub-task, and also the effects of stimulus dimensionality. The categories have been ordered in Figure 4.16 in terms of the hypothesized differences in strategy that would be employed by the two groups. As predicted, there are substantial differences; the bar-chart group are clearly taking a more analytic approach in forming clusters whilst the 'scenic' strategy, primarily adopted by the faces group, can be interpreted as a placement in terms of some 'distance' from the referent item. Thus H1 is confirmed in terms of the behavioural protocols. There is also an ordered interaction with the degree of stimulus dimensionality, in that the differences in strategy become more pronounced with increasing dimensionality. This is as predicted by H2.

A dimensional analysis of the reference ranking data was made using the ALSCAL program, with the initial configuration 'fixed' on the constructed values of the stimuli sets. Figure 4.16 shows the dimensional weightings of the two groups for the two, three and four dimensional stimuli. The dimensional weightings in this case, can be taken as an indication of the accuracy of judgement on a particular dimension. There is an obvious

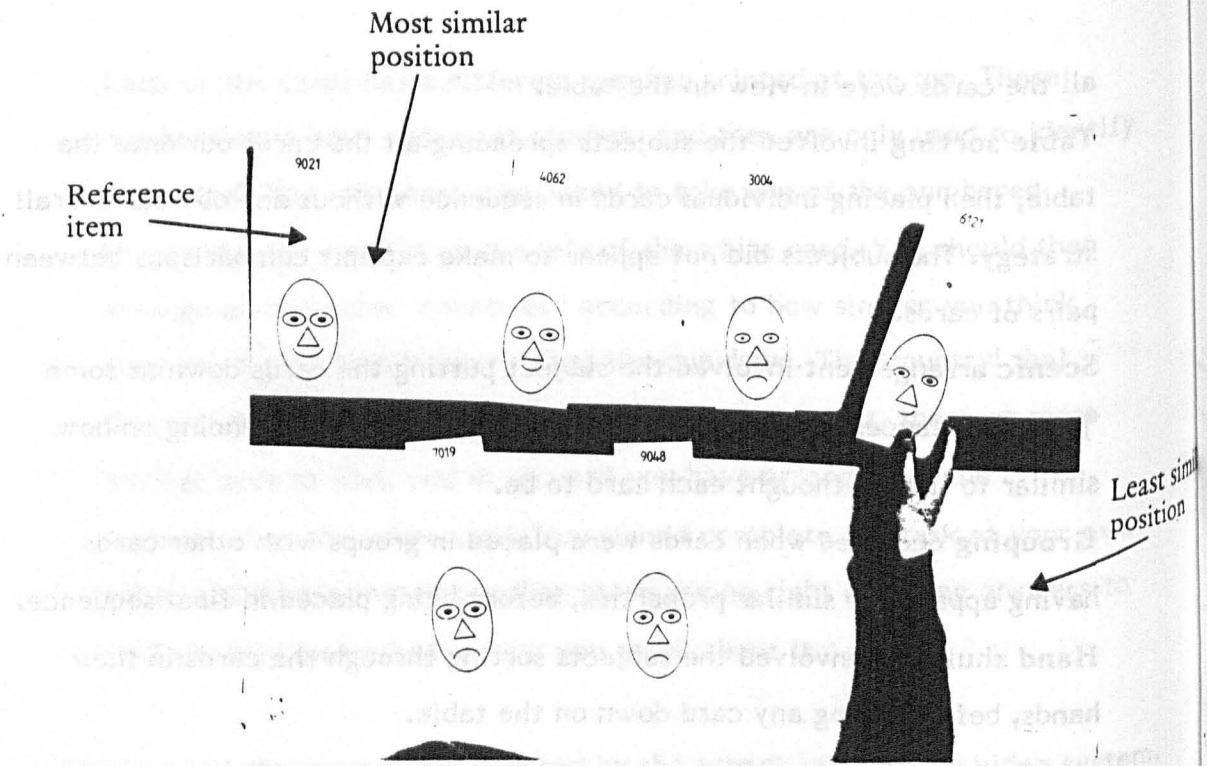


Figure 4.11 A 'scenic' strategy being used with the cartoon face format

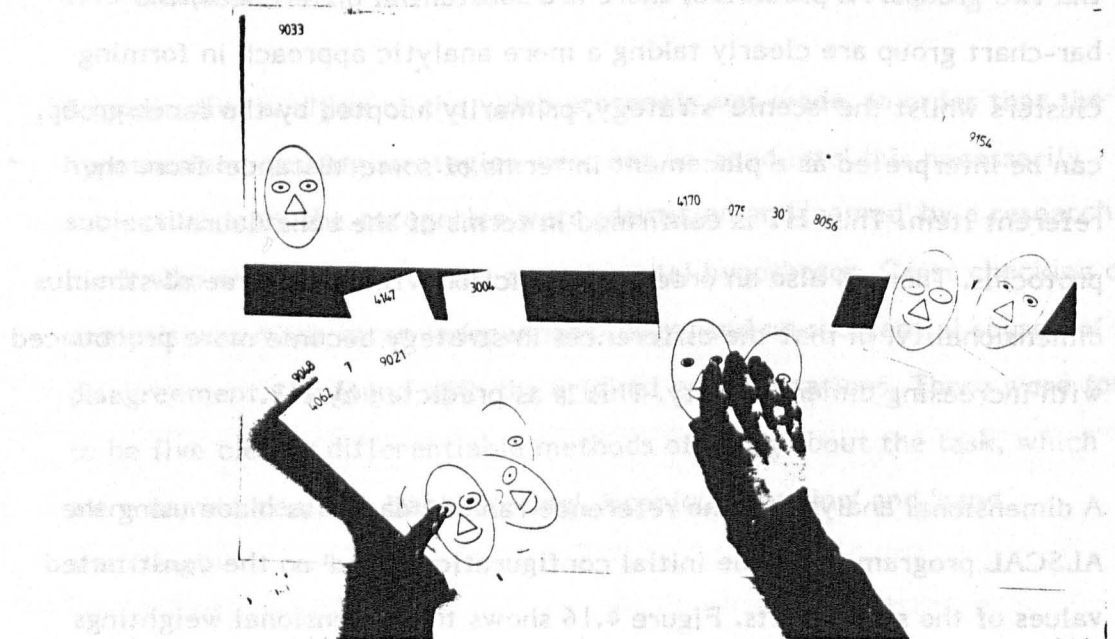


Figure 4.12 A 'groups' strategy being used with the cartoon face format

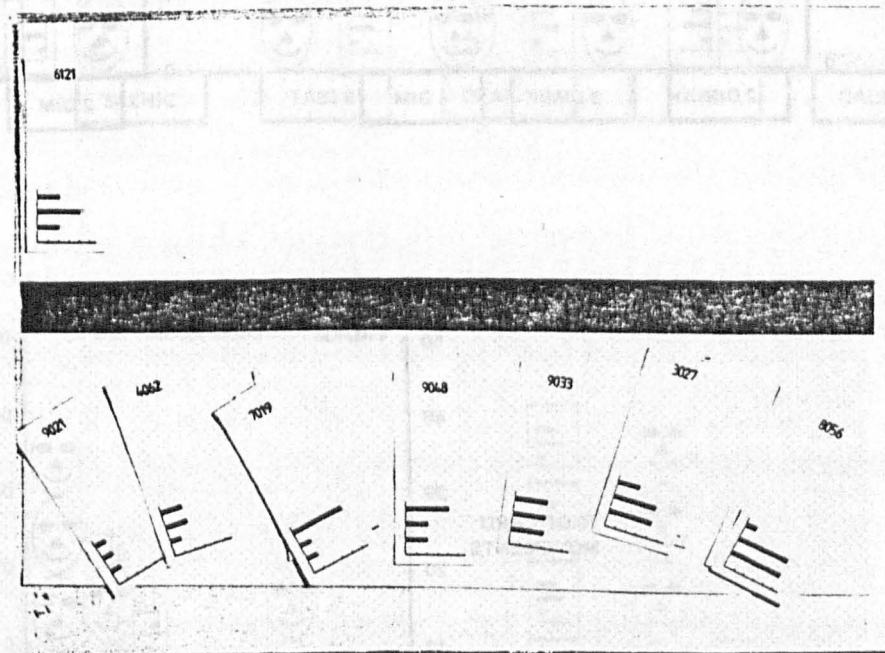
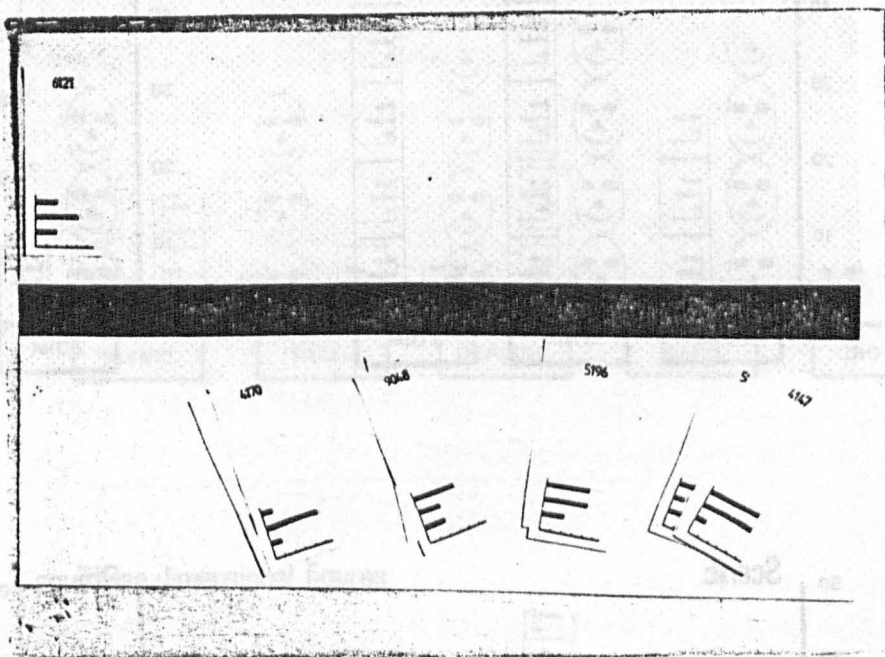


Figure 4.13 Two stages in the use of a 'groups' strategy with the bar charts

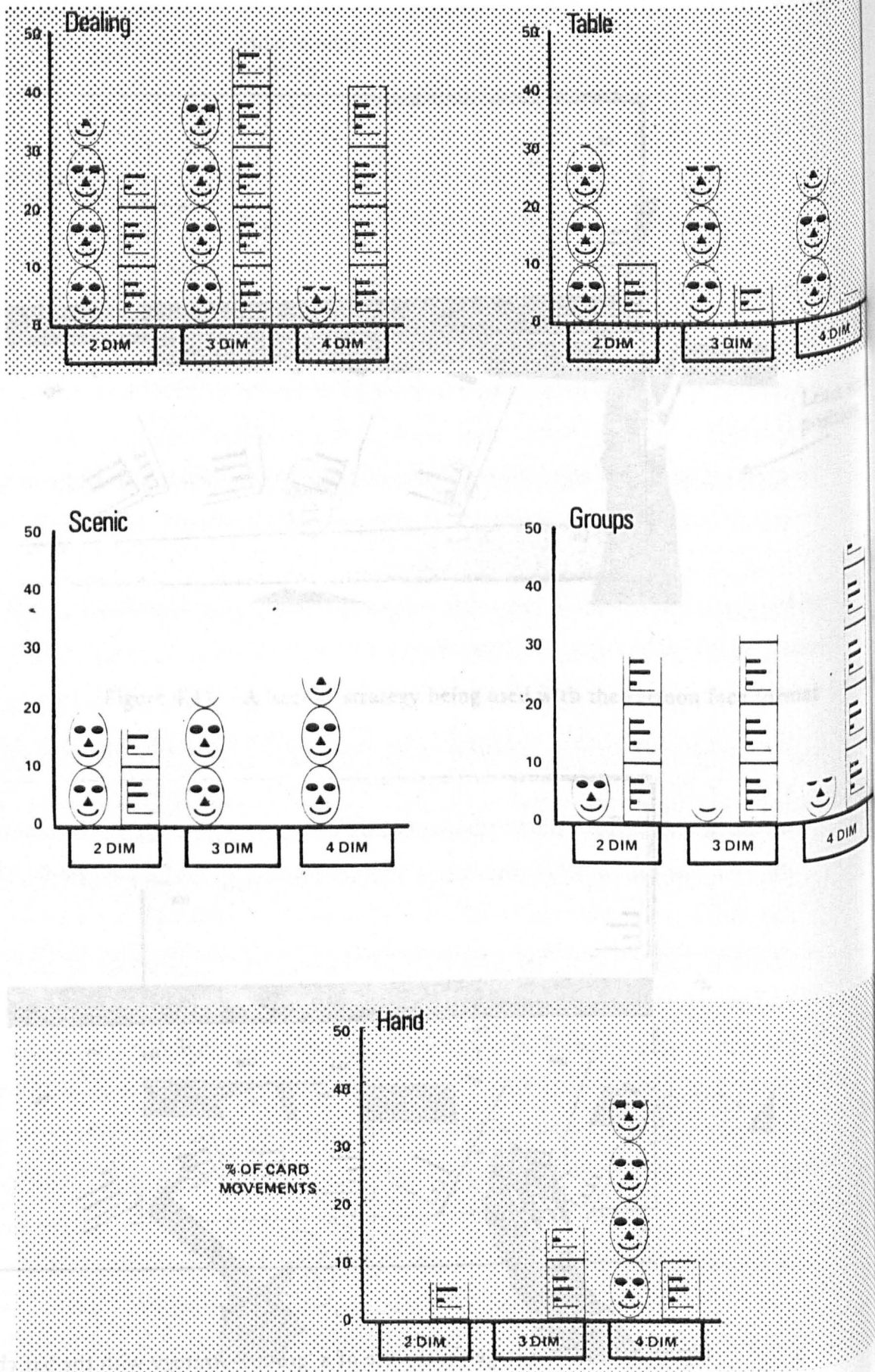


Figure 4.14 Categorization analysis of ranking strategies: by category, and by subject group

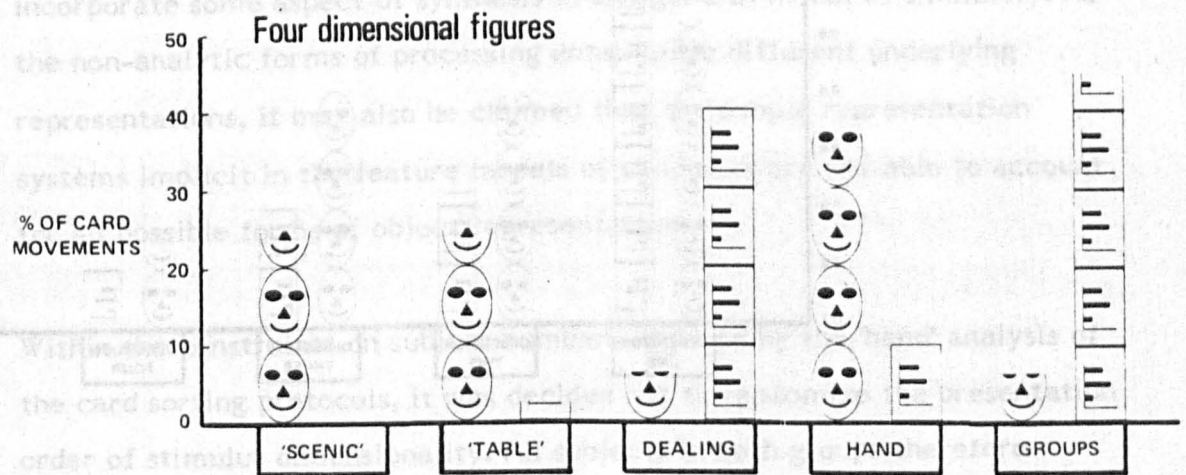
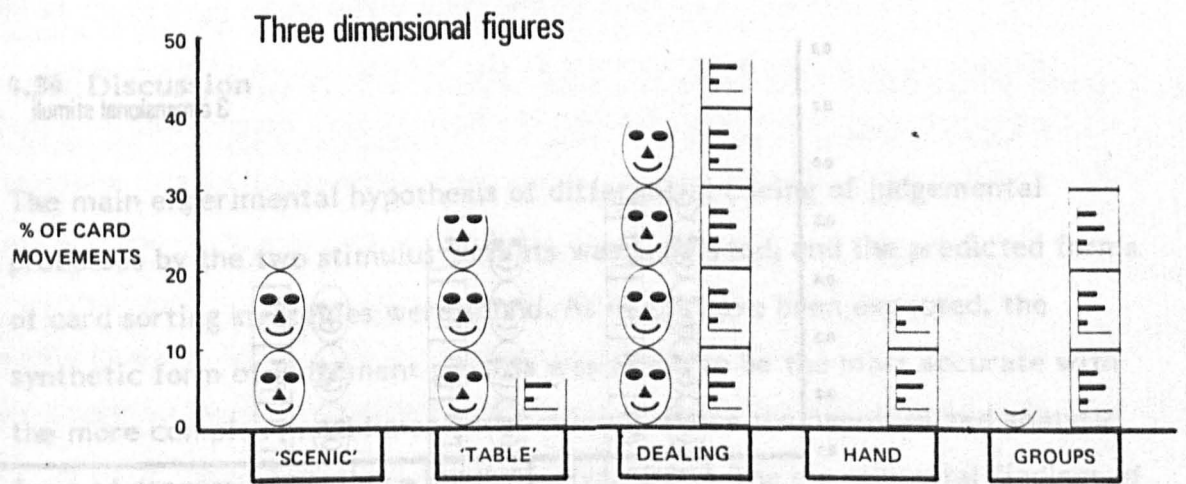
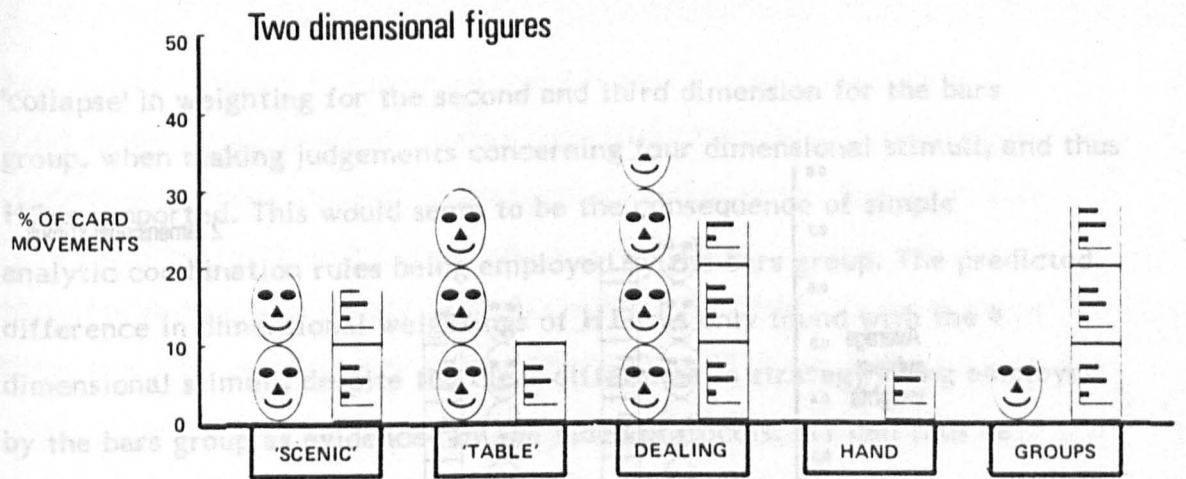


Figure 4.15 Categorization analysis of ranking strategies: by dimensionality, and by subject group

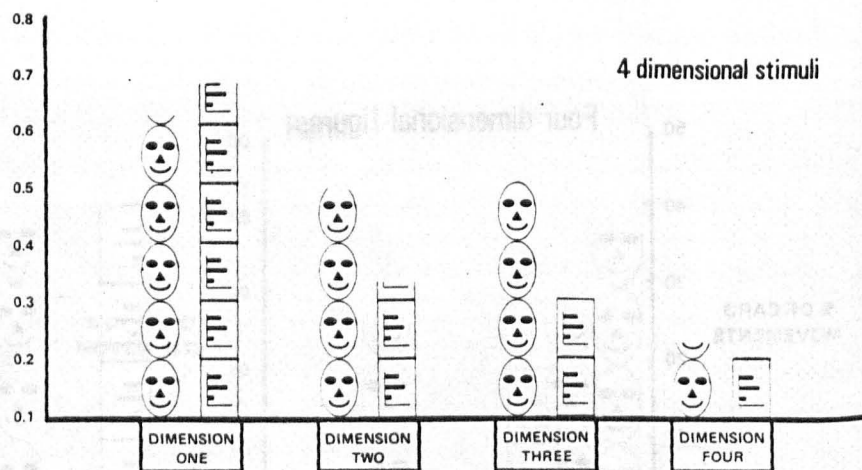
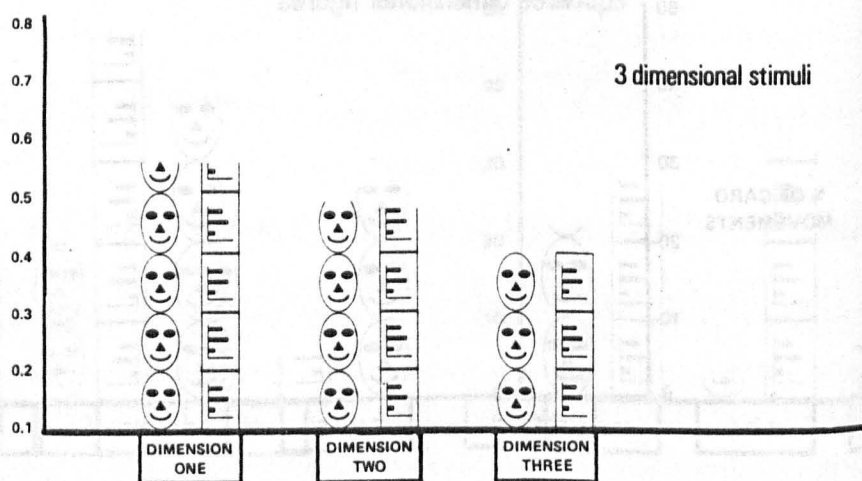
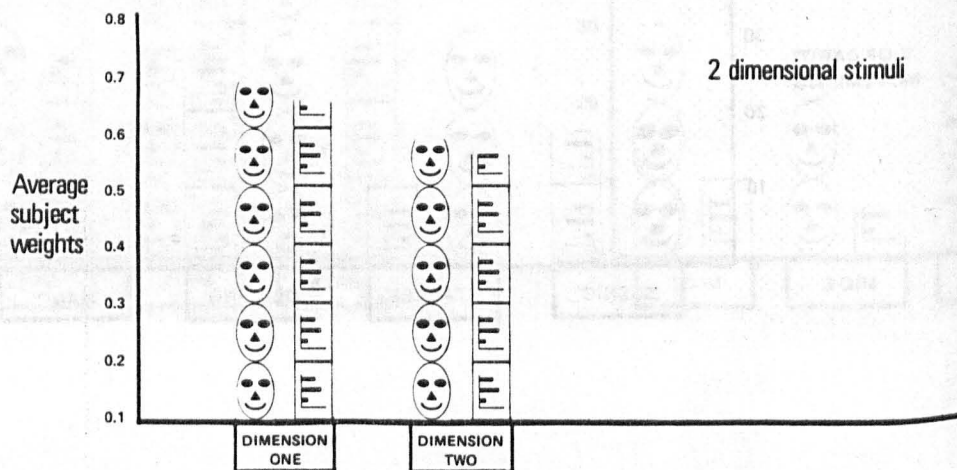


Figure 4.16 Dimensional weightings for 2, 3 and 4 dimensional stimuli: by subject group

'collapse' in weighting for the second and third dimension for the bars group, when making judgements concerning four dimensional stimuli, and thus H2 is supported. This would seem to be the consequence of simple analytic combination rules being employed by the bars group. The predicted difference in dimensional weightings of H1 was only found with the 4 dimensional stimuli, despite the clear difference in strategy being employed by the bars group as evidenced by the video protocols. H1 can thus be considered to be only partially supported by this analysis of the similarities data.

4.34 Discussion

The main experimental hypothesis of differential cueing of judgemental processes by the two stimulus formats was supported, and the predicted forms of card sorting strategies were found. As might have been expected, the synthetic form of judgement process was found to be the most accurate with the more complex multidimensional stimuli; where the hypothesized analytic form of processing would be most disadvantaged. The experimental findings of the differences in strategy employed by subjects whilst making similarity judgements, further supports our contention concerning the necessity to incorporate some aspect of synthesis in any general model of similarity. As the non-analytic forms of processing entail quite different underlying representations, it may also be claimed that the simple representation systems implicit in the feature models of similarity are not able to account for all possible forms of object representation.

Within the constraints on subject numbers imposed by the 'hand' analysis of the card sorting protocols, it was decided not to randomize the presentation order of stimulus dimensionality. All subjects in both groups therefore received the appropriate format in order of increasing dimensionality. This would appear to leave the experiment open to the suggestion that the results might be the consequence of some order effect. However the claim made for

this experiment was only that the two encodings of the information matrix would lead to different cognitive processes. Any order, practice or fatigue effects brought about solely by the manipulation of information format or matrix dimensionality would not weaken this claim. Their differential presence would count against the claims of the feature models in that they do not possess any explanatory representation or process for them, whilst the proposed model does.

The prediction was made in H1 of better performance, in terms of dimensional recovery for the faces group. That this is not found, except in the four dimensional case, despite the clear differences in strategy being employed, is probably an example of the 'robust' nature of the scaling algorithms being a mixed blessing. We have already commented on a major problem for studies attempting to compare hierarchical cluster analyses with dimensional scaling analyses; namely that the two sets of algorithms are optimizing on quite different aspects of the data, and will therefore be differentially affected by the same pattern of data error. It would appear that for low dimensionalities, the forms of error introduced by the 'Groups' strategy appears to be of the same order to the scaling algorithms as normal judgement errors within the 'Scenic' strategy. The 'power' of the scaling algorithms was thus found to be such that significant differences in adopted strategy by subjects may be hidden, and this has important implications for studies dependent upon the elicitation of estimates of similarity. As a result of stimulus format and complexity, subjects may be taking quite different perspectives on similarity judgement tasks, both with respect to different forms of task, and also with the assumptions of the experimenter. Except with high (relative to that used in most experimental situations) dimensionality stimuli, this may not become apparent using conventional analyses.

CHAPTER FIVE **Applied studies**

5.1 Overview

The importance of a theory providing a useful model of similarity judgement, other than within tightly controlled experimental paradigms was argued in Section 1.4. Three studies are detailed here in which the previously developed experimental methodology is applied to 'real-world' problems. The rationale for the inclusion of these studies is to demonstrate both the applicability of, and the necessity for, a synthetic component in any model of perceived similarity. The studies represent areas of investigation where attempts to elicit information in a manner requiring an analytic form of judgement is shown to be more or less inadequate. In each case, the comparative success of methods requiring subjects to make synthetic relational judgements is demonstrated; in terms of eliciting information useful to the purpose of the applied investigation.

The first study assumes, but also seeks to facilitate, a dimensional representation of the problem domain. The intention being to study group differences in perceived relations within such a framework. It is concerned with the perception of pain following different medical treatments for severe burns, and different forms of surgical operation. This study is important to our argument, in that it represents an attempt to elicit similarity judgements in an operative knowledge domain, where the effects of language as a mediational representation are not dominant. In the second study, the development of an appropriate 'meta-language' by students taking a microprocessor course is investigated. It is shown that even novice programmers are capable of perceiving relations between important computing concepts in other than an analytic fashion. The importance of this finding is discussed in terms of the possible differentiation between mistakes as failures of knowledge, and mistakes as a mistaking of the problem. The third study deals with the appreciation of possibly contentious arguments in

texts, and represents an attempt to 'tap' individuals' judgements of concept relations in a knowledge domain where the definition of concepts is not axiomatic.

The discussion relating to each applied study is made primarily in terms of the relative efficacy of eliciting a dimensional representation of the relational aspects of the particular objects, as against the use of a corresponding analytic judgement task. A discussion of the significance of these findings in terms of our general argument is deferred to the final section of this chapter.

5.2 The perception of pain

5.2.1 Rationale

The majority of the studies detailed in Chapter 2 have been to a great extent dependent upon the mediating role of language, which may provide an analytic framework for the decomposition of the stimulus items. It was suggested in Section 2.22 that the 'feature' theories concerned with the estimation of relational similarity have concentrated on the easily mathematised, analytic aspects of the judgemental task, to the exclusion of its possible synthetic aspects. It is claimed that the latter would allow the individual to create a general framework, within which they may make some personal sense of the tasks required of them. This first applied study illustrates the advantages of a synthetic, or dimensional, approach within certain domains. It is intended to exemplify the usefulness of effectively imposing a holistic framework upon the individual required to make particular relational judgements. The study detailed here was made in two parts, the first was concerned with the subjective perception of post-operative pain following two alternative treatments for severe, but local, burns, and the second with two different forms of incision in upper abdominal operations.

Clinical evidence regarding the relative speed of healing resulting from the use of the different treatments would lead one to expect clear group differences in perceived pain in the first part of the study. The more recently developed treatment entails covering the burn with a thin plastic film, alleviating the necessity for frequent dressing changes. The healing time for the 'bandage' group has been shown to be significantly longer than for the 'film' group (Neal, Whalley et al, 1981). No clear differences between the various forms of incision, in terms of pain and perception, are reported in the clinical literature. However a horizontal incision might be expected to be more traumatic than a vertical one, in that it involves cutting across a greater number of nerve fibres, and was thus expected to cause greater pain with bodily movements for a longer period. In the first part of the study it was intended to contrast the relational comparison method with the alternative questionnaire technique described below. In the second part it was intended to show that even with the most intense form of pain, the relational method would still provide an interpretable dimensional structure within which the hypothesized group differences could be made apparent.

Numerous researchers in this area have commented on the difficulties experienced in eliciting useful linguistic descriptions of experienced pain, whether for individual diagnostic feedback or for exploring group differences. Although some researchers (e.g. Huskisson, 1974) have attempted to obtain analogue comparison measures between the pain currently experienced and some previous state, the dominant methodology employed at present involves checking words from a lengthy questionnaire. In the context of our argument, this represents an attempt to force the patient to analytically decompose their experienced pain. Although some success with the questionnaire technique has been reported in the literature (e.g. Melzack, 1975), it has generally been found to only be of use with individual patients; who effectively have to be trained how to respond with it.

The technique employed in the present study required the subjects to relate the pain they were feeling to a set of commonly occurring painful events. It would be predicted from the relational model developed in Chapter 3 that this more holistic form of comparison would both be more meaningful to the patients, and also cause them to make more accurate judgements concerning their perceived pain. Our argument then is that this elicitation method should result in a common stable dimensional structure which would enable group differences in subjectively perceived pain to be made apparent.

5.22 Method

Task The reference ranking similarity elicitation task was used in both the 'burns' and the 'incision' trials. Because only a small number of stimuli were employed, and there were no time constraints, complete data sets were obtained, i.e. each of the stimuli were used as the reference item in turn. The common painful experiences used as the stimulus items in each trial are listed in Appendix C. These were all signified by single words or short phrases e.g. 'headache', 'tooth filled', 'your pain' etc, and were typed onto small index cards for use in the ranking task. After the reference ranking task the patients in the 'burns' trials were also required to complete the Melzack Pain Questionnaire (also contained in Appendix C). Both parts of this study employ a 'between-subjects' design.

Subjects and sessions. The eighteen subjects in the 'burns' trial were all outpatients at the Leeds General Infirmary, and were randomly assigned to either the control group receiving the conventional gauze burns dressing, or to the experimental group receiving a newer 'plastic film' treatment. The reference ranking task and the questionnaires were administered one week after each patient's treatment had commenced. The various tasks were individually administered whilst the patients were waiting for treatment, and no payment was made for completing the forms. The twelve subjects in the 'incision' trial were all patients at the Leeds General Infirmary undergoing major upper abdominal operations. Patients were randomly assigned to the

groups to receive either the vertical or the horizontal operational incisions. The reference ranking task was individually administered to each patient one week after their operation, and again no payment was made.

5.23 Results

ALSCAL analyses were made, and a two dimensional plot of the group space of the stimulus items for the burns trial is shown in Figure 5.1. It is possible to interpret the dimensions as being related to the 'intensity' and 'duration' of the painful experiences. The individual patients' dimensional weightings are shown in Figure 5.2, a clear difference between the two groups of patients being apparent. As predicted, the 'bandage' group show a markedly higher weighting for the 'duration' dimension. The Melzack questionnaire has been developed around various factor analysis studies, and for analysis, the groups of words are grouped into three categories of 'sensory', 'effective' and evaluative terms. Comparative analysis of the data for the two burns groups revealed no significant group differences for any of the factor categories. Analyses were made using the Student t test:

	Film	Bandage	
Sensory			
Mean	5.83	5.33	$t = .36, p (-3.60 \leq \mu \leq 2.60) = .95$
SD	1.72	2.94	
Affective			
Mean	2.17	2.33	$t = .19, p (-1.75 \leq \mu \leq 2.08) = .95$
SD	1.72	1.21	
Evaluative			
Mean	2.67	2.33	$t = .62, p (-1.50 \leq \mu \leq 0.86) = .95$
SD	1.03	0.82	

Table 5.1 Comparative analyses of the questionnaire data in the burns study

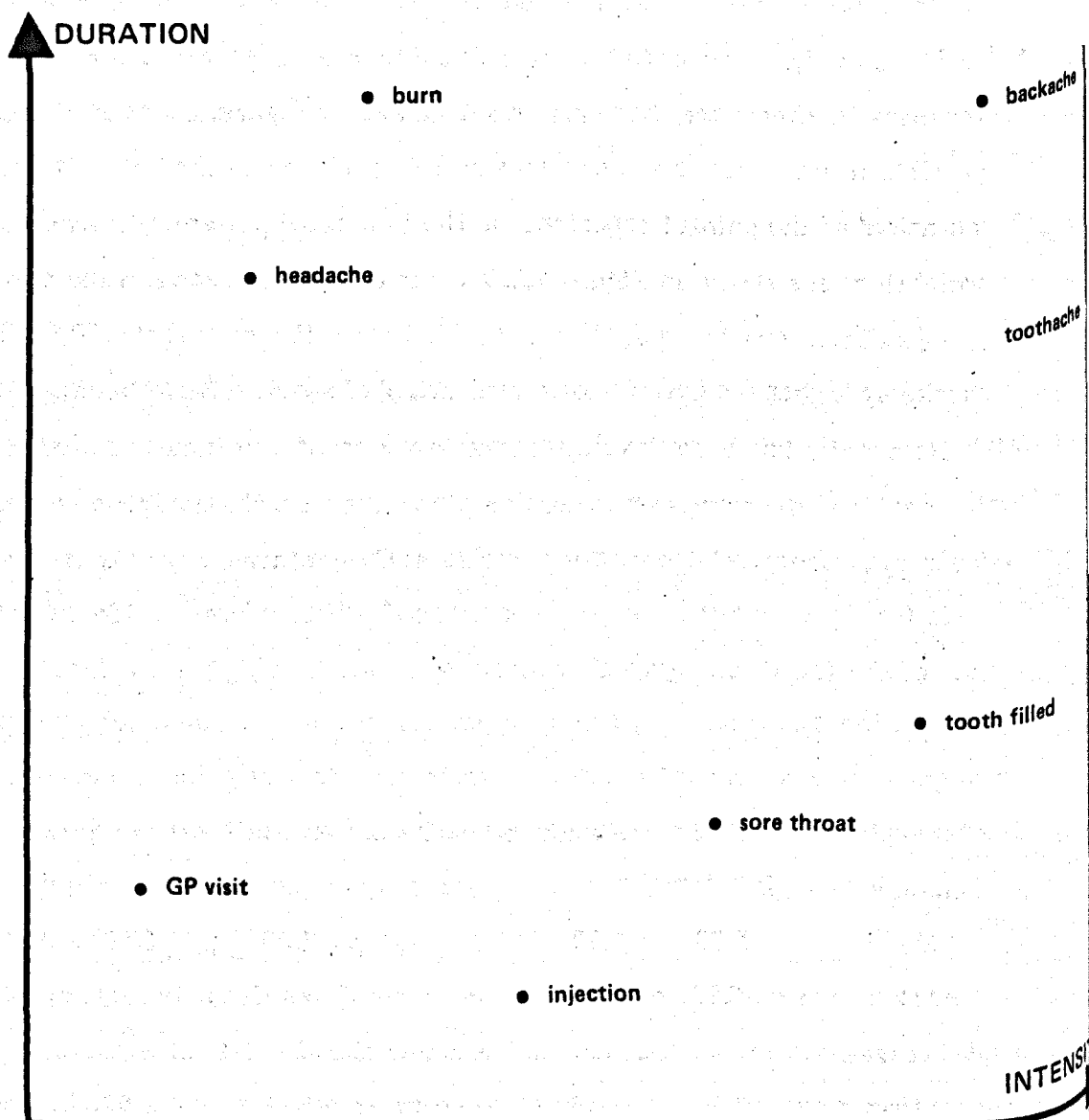


Figure 5.1 Two dimensional analysis — burns patients (ALSCAL-4 individual differences analysis, 18 patients, 8 stimuli, STRESS = .247, RSQ = .549)

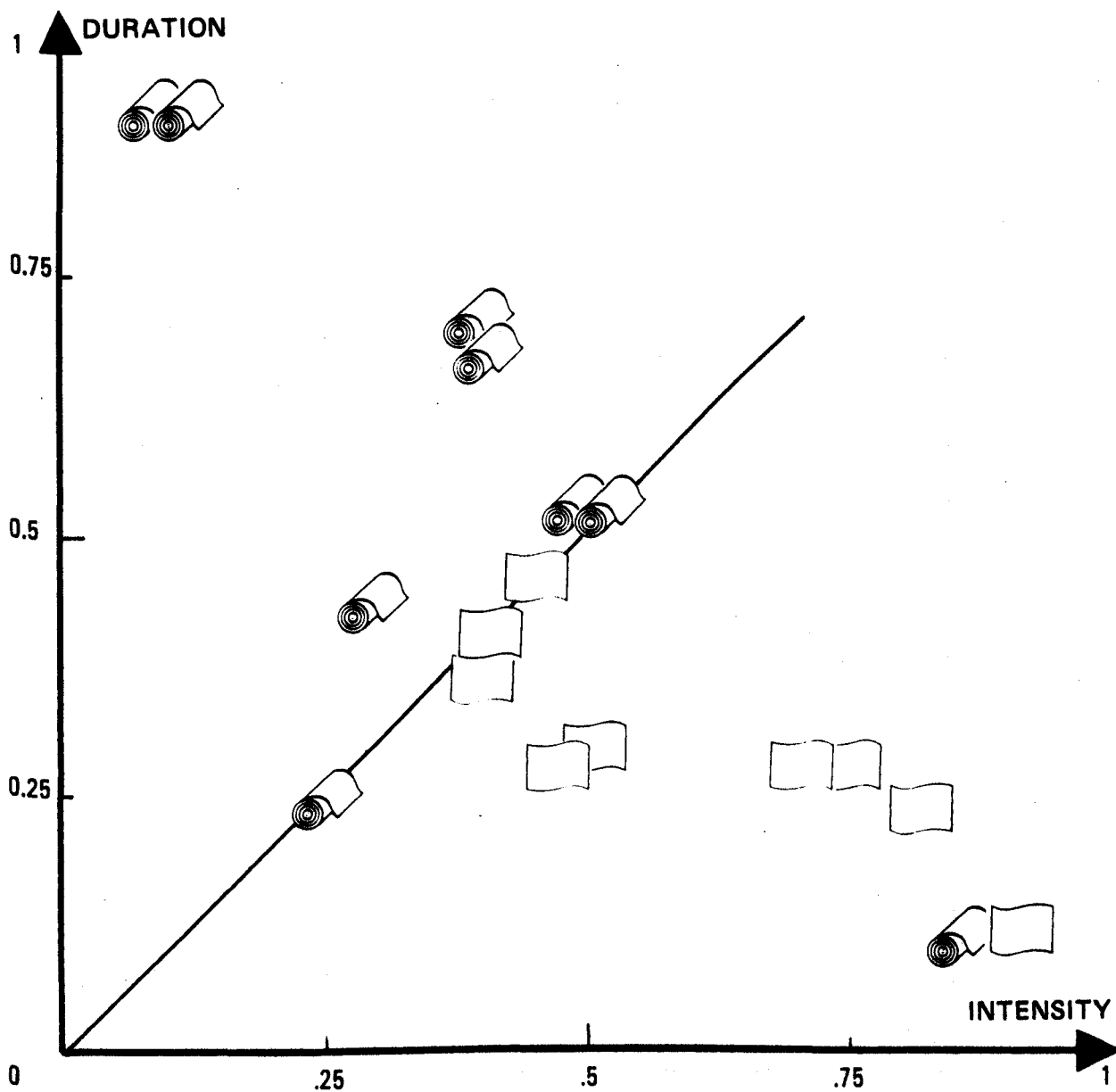


Figure 5.2 Subject weights for the bandage and plastic film burns treatments
(ALSCAL-4 individual differences analysis, 18 patients, 8 stimuli, STRESS = .247
RSQ = .549)

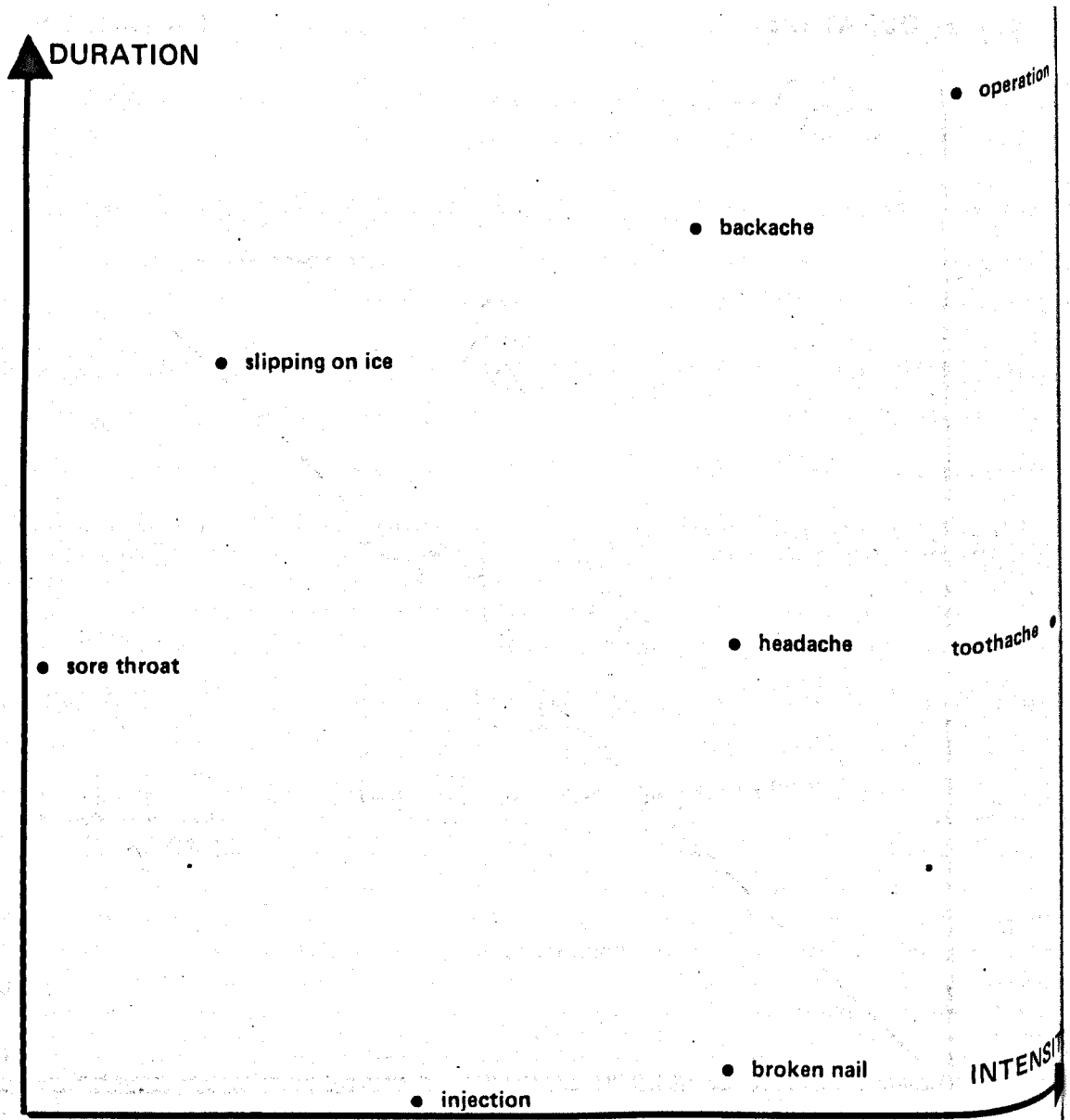


Figure 5.3 Two dimensional analysis — vertical and horizontal incision patients
 (ALSCAL—4 individual differences analysis, 12 patients, 8 stimuli, STRESS = .188
 RSQ = .841)

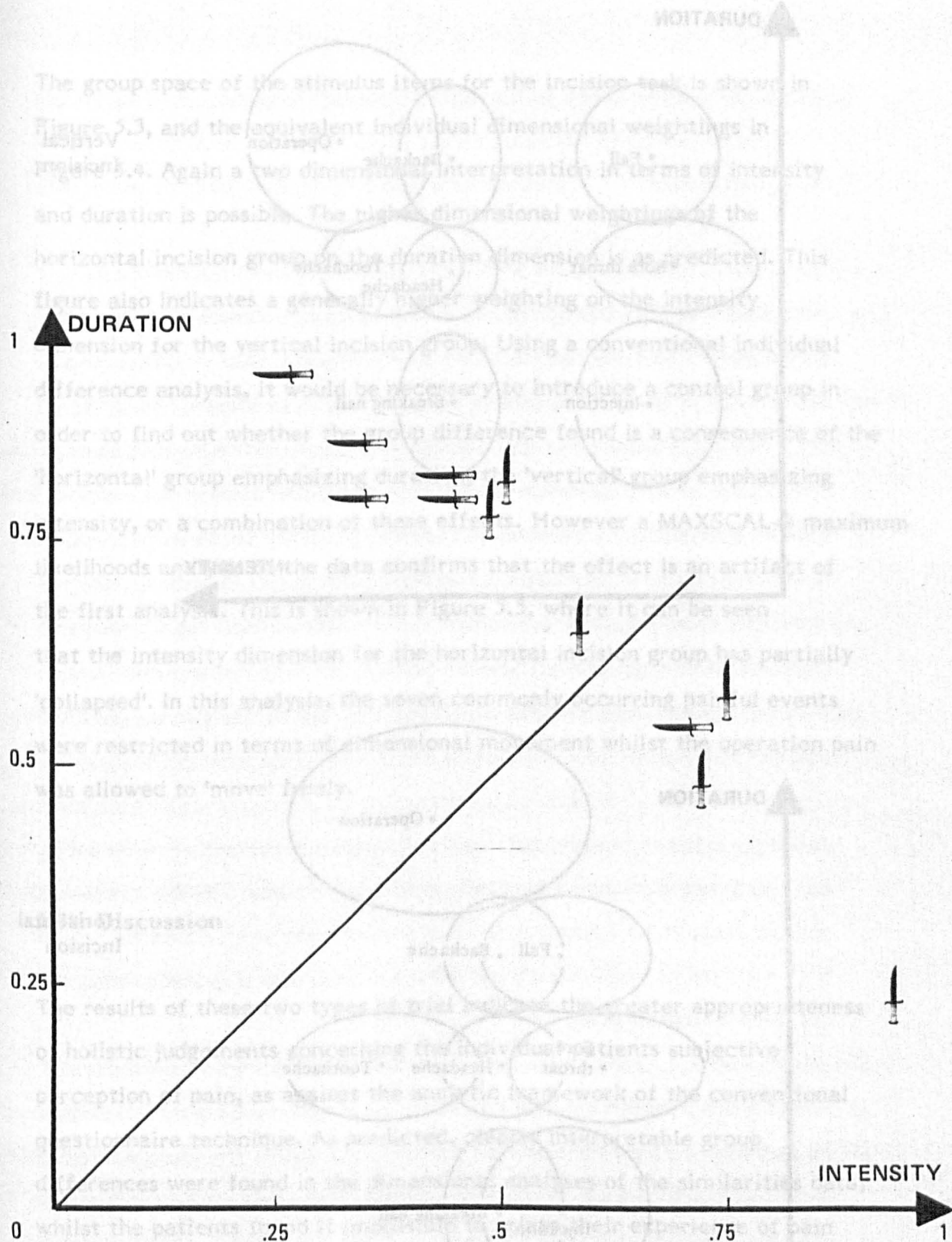


Figure 5.4 Subject weights after operations involving vertical and horizontal incisions (ALSCAL-4 individual differences analysis, 12 patients, 8 stimuli, STRESS = .188, RSQ = .841)

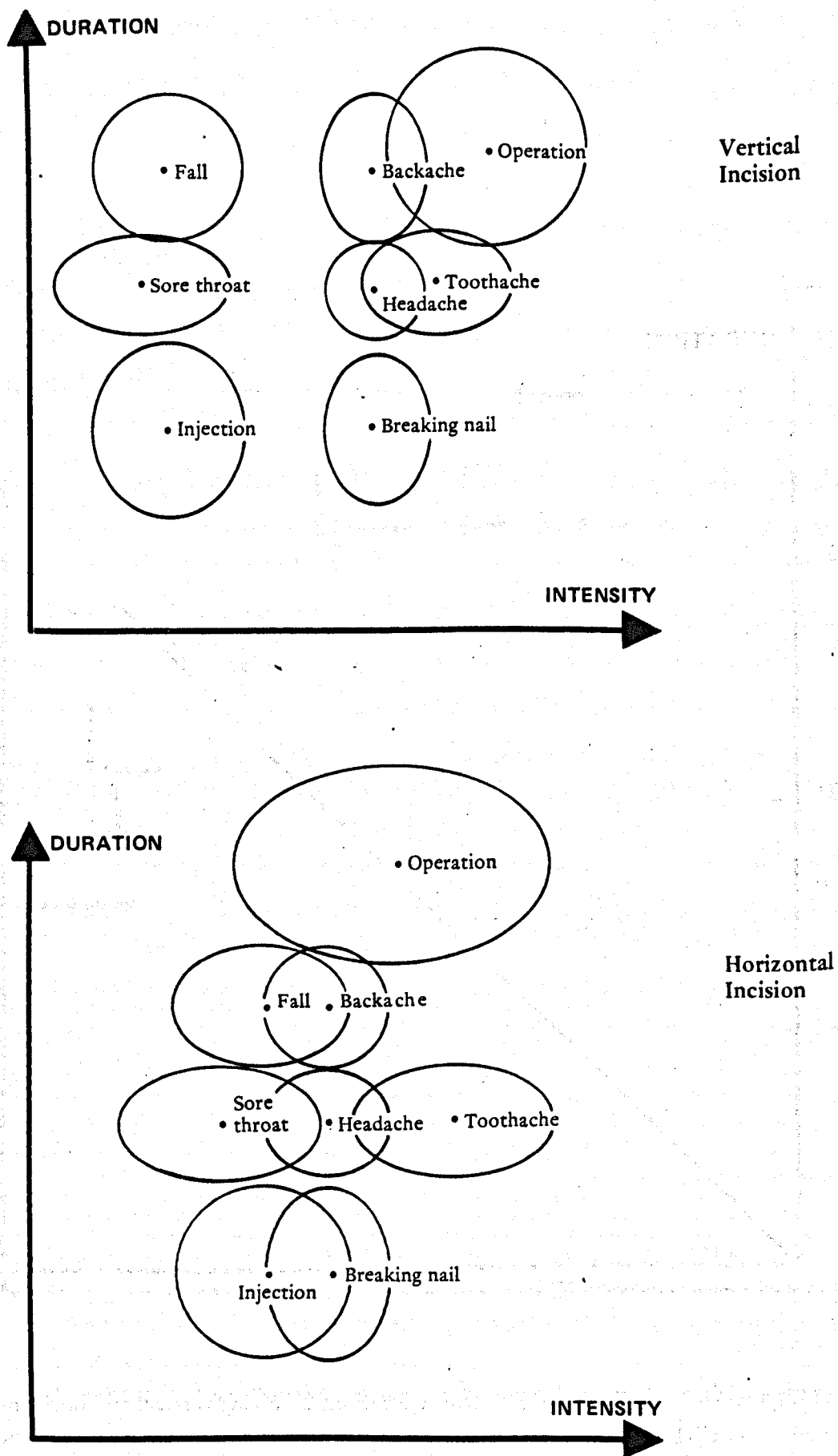


Figure 5.5 MAXSCAL-4.1 analyses of the Incision data (ellipses indicate 95% confidence regions)

The group space of the stimulus items for the incision task is shown in Figure 5.3, and the equivalent individual dimensional weightings in Figure 5.4. Again a two dimensional interpretation in terms of intensity and duration is possible. The higher dimensional weightings of the horizontal incision group on the duration dimension is as predicted. This figure also indicates a generally higher weighting on the intensity dimension for the vertical incision group. Using a conventional individual difference analysis, it would be necessary to introduce a control group in order to find out whether the group difference found is a consequence of the 'horizontal' group emphasizing duration, the 'vertical' group emphasizing intensity, or a combination of these effects. However a MAXSCAL-4 maximum likelihoods analysis of the data confirms that the effect is an artifact of the first analysis. This is shown in Figure 5.5, where it can be seen that the intensity dimension for the horizontal incision group has partially 'collapsed'. In this analysis, the seven commonly occurring painful events were restricted in terms of dimensional movement whilst the operation pain was allowed to 'move' freely.

5.24 Discussion

The results of these two types of trial indicate the greater appropriateness of holistic judgements concerning the individual patients subjective perception of pain, as against the analytic framework of the conventional questionnaire technique. As predicted, clearly interpretable group differences were found in the dimensional analyses of the similarities data, whilst the patients found it impossible to relate their experience of pain to the analytic terms of the questionnaire. Any general model of similarity judgements must be capable of making useful predictions in such extra-linguistic domains, and these results must further count against the claims of the feature analytic models with regards to generality.

5.3 Expert knowledge in a micro-processor course

5.3.1 Rationale

The mental representation of the concepts involved in a problem is recognized as influencing the individual's perception of a problem; and consequently how they might solve it. Shavelson (1972) showed how as students progressed through a physics course, their knowledge of the structural relationships of the discipline became more like that of experts. This study, as with others in the domain, made use of both dimensional and cluster analyses of individual's perceived similarity between concepts. In a review of methods for mapping cognitive structures, Preece writes that "problem-solving tests are, perhaps, too drastic as tools for the initial exploration of cognitive structure -" (1976, p1). The various elicitation methodologies "- tamper less with what they are seeking to measure". One of the earliest studies concerned with eliciting knowledge representations, Wainer and Kaye (1974), used a dimensional representation of the students' concept relations in order to examine changes in the perceived importance of central topics in an undergraduate psychology course. However, as we shall note in Section 5.4.1, most elicitation studies have been concerned solely with domains of 'hard' knowledge where a dimensional representation has appeared inappropriate.

As **problem perception** has now become recognized as a crucial component of problem solving performance, so investigations of the change in problem perception with the acquisition of expertise has also received increasing attention. Schoenfeld and Herman (1982) has extended Shavelson's work to show how students' perceptions of the structure of mathematical problems changed as a consequence of a course of study. Kahney (1982) has also indicated that novice programmers, in a high level AI language, may quite quickly develop the structural concept relationships of an 'expert', but not necessarily be able to tackle problems in the same way. The 'novice-expert' distinction is thus shown to involve more than a difference in a surface

expertise with the terms of a topic. The nature and individual perceptions of the problems to be solved using a particular set of concepts, or computing language, must be taken into account.

The intention of this study is to evaluate the importance of the development of appropriate meta-languages, within which the student may understand and solve a problem, e.g. the appreciation of the 'top-down' approach to the functional decomposition of computing problems in terms of 'data-flow'. Our general argument has been that in certain contexts individuals can and do think in more general terms than the feature models of similarity would imply or allow; in terms of 'synthesis' rather than 'analysis'. The acquisition of such an abstract representational ability might provide a meta-language within which the operation of even highly abstract relations may be discussed. Cognitive structure theorists such as Schroeder et al (1967) have suggested that abstract functioning in many cognitive domains would be dependent upon the development and use of superordinate dimensions. Such dimensions are not immediately related to the specific subject matter of the domain, but form a meta-language which may facilitate discourse about the sub-domain. The possibility is examined here of using the individual's facility in representing concepts dimensionally as a criterion for distinguishing operational from 'expert' knowledge.

The study was intended to determine whether an individual's development of such meta-languages could be examined using data which was derived from their perception of similarity relations between items in the sub-domain. It was hoped that the extra-linguistic, and less intrusive, form of such a task would be useful in determining the nature of novice-expert differences. These differences might well be masked in the alternative, feature dependent, memory or clustering tasks currently made use of in the literature. The prediction of our general argument, is that it would be possible, and useful, to elicit a meaningful dimensional representation of the course concepts from students, by biasing them towards synthetic rather than analytic judgements.

5.32 Method

A pilot study was first conducted using eight computer 'professionals', mainly drawn from the Technology Faculty of the Open University. The intention was to check whether the elicitation methodology was viable, and to provide for the analysis of a stable 'prior' representation which would facilitate the later analysis of the students' data. The subjects were given thirteen small index cards each presenting a key word or phrase related to an introductory microprocessor course. Eleven of the words were derived from the detailed experimental course material. The two more abstract ideas of 'flow' and 'control' were added as they were discussed in other course material, and appeared likely to suggest salient dimensions to the students. Details of the stimulus material are contained in Appendix D. After reading through the cards for a few minutes the subjects were then requested to make similarity judgements between the items using the reference-ranking technique. Whilst making the judgements, they were asked to think of the items in terms of their use in the microprocessor course; of which they were all knowledgeable.

The novice-expert study was conducted using Summer School students engaged in the Open University foundation technology course, T101. As part of the summer school the students took a one day course which involved the use of a small microprocessor kit and a self-instructional text. [The study detailed here was in fact part of a wider investigation of the use of 'help' algorithms in the text using a study protocol recorder to monitor the pages of the text being used, and also the key depressions being made on the micro kit, cf Whalley, 1982.] In the last half hour of the one day course, twelve students out of the class of thirty, were asked to complete the reference ranking task. Out of the thirteen stimuli, four were chosen as reference items, and each subject received the four reference items in a random sequence. The students had all completed the problems set for them, and had thus experienced fewer problems than the rest of the class. Out of the twelve students, four had not been able to start the ranking task with the

first reference item within five minutes, and were switched to a rating task employing the same 13 items. [This questionnaire was also completed by the remainder of the students, although its analysis is not discussed here.] The 'data' from these students is not included in the analysis reported. Two tutors on the course and the postgraduate student acting as demonstrator also completed the reference ranking task.

5.33 Results

The group space derived from an ALSCAL analysis of the prior study data is shown in Figure 5.6. The data fit the intuitively obvious dimensions for any computer language of 'flow-control vs. location-value' and 'active vs. passive'. It is not clear from this analysis whether the structure truly represents a 'folded' dimension, or whether it is an artifact (cf Takane and Carroll, 1982). However it would make some sense to consider the structure as a folding of the dominant 'flow-value' dimension by the more abstract 'active-passive' one. An individual differences analysis of the Summer school data set was made, using ALSCAL, 'fixed' on this prior configuration. The resultant dimensional weightings are shown in Figure 5.7 and it can be seen that the individual weightings are fairly widely distributed. It is useful to note the position of the two tutors who completed the task, as they may be seen to represent some 'true' value concerning the general orientation of the course.

The prior study had indicated that subjects with some experience of high level computing languages (e.g. BASIC, but particularly PASCAL) weighted most heavily the 'flow-value' dimension. However this course was primarily intended to demonstrate the operation of the accumulator and data registers, as is indicated by the relative weighting of the tutors. One of the more interesting aspects of this study, is that it provided independent confirmation of a misperception of the course's purpose by three of the four students who had had prior experience of a high level language. These

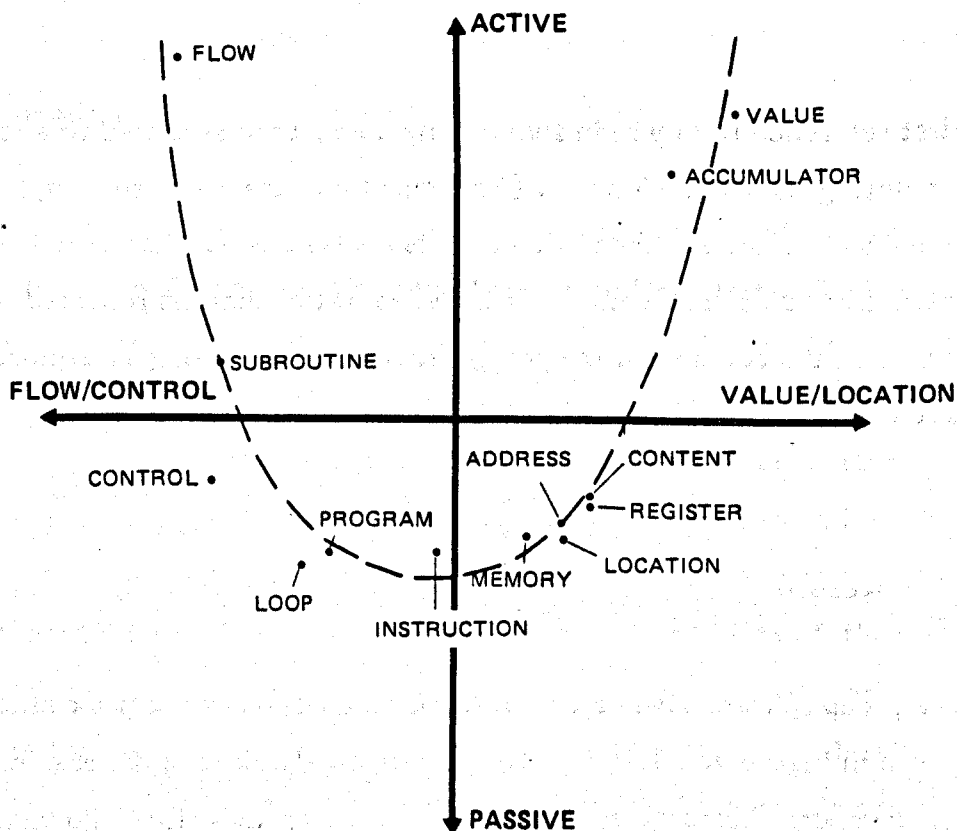


Figure 5.6 Two dimensional analysis of the professional programmers' judgements of the micro-processor course concepts

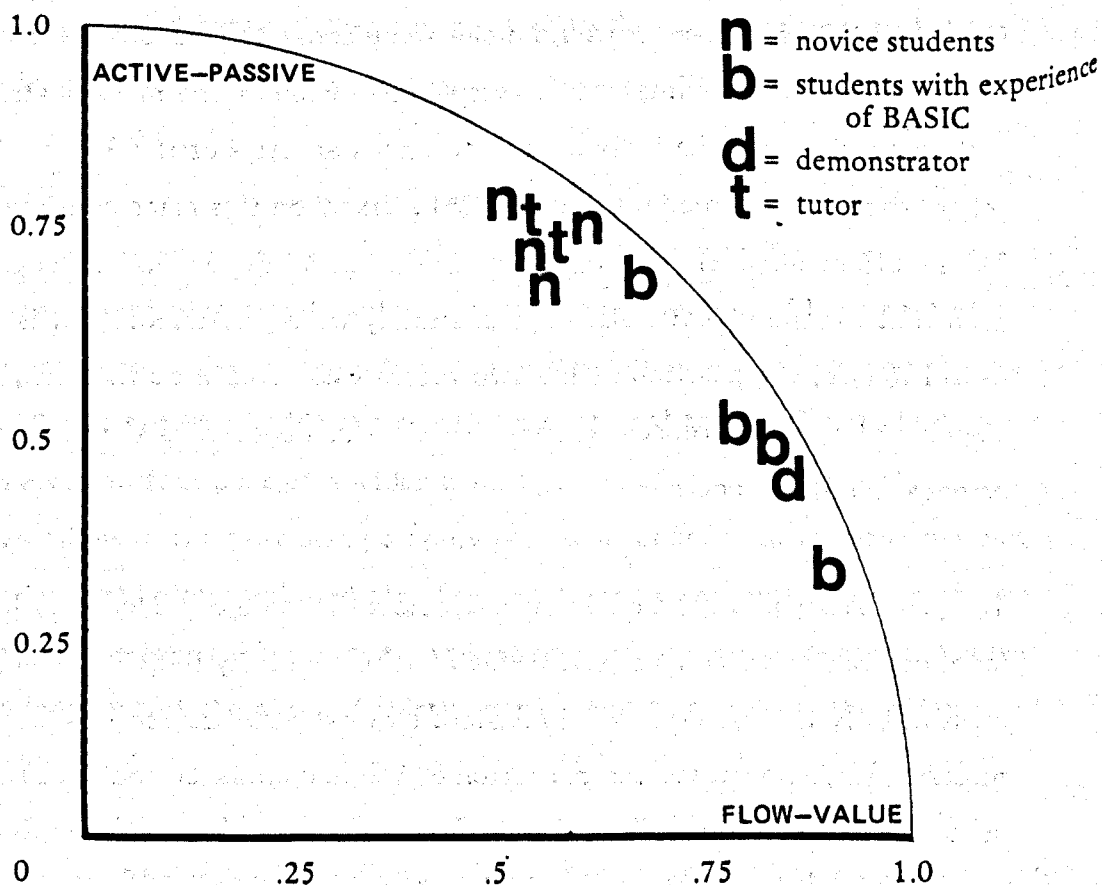


Figure 5.7 Dimensional weights of students and tutors (ALSCAL-4 analysis, 'fixed' on the configuration of Figure 5.6, 13 stimuli, 11 subjects, SSTRESS = .133)

students are in the group emphasizing the 'flow-value' dimension, along with the demonstrator who also had only had experience of high level programming. The protocol recordings showed that these students effectively skipped the early sections of the course, in order to move on to the later flow and control experiments. However, although they knew far more about computing in general than the others, they continued to experience difficulties over detailed issues and had to turn back to the early section more often.

5.34 Discussion

The results of this study indicate that a majority of the 'novice' students were capable of forming a dimensional representation of the topic concepts. Examination of the individual differences found with respect to these dimensions suggests that they may be interpreted as a valid meta-language for this topic domain.

The evaluation of dimensional structures such as those found here might aid the study of meta-language development. Such an evaluation would obviously need to be complemented by an analysis of the learner's verbal protocols, and a deeper study of the pedagogic factors which might facilitate the development of such reflective reasoning. Following Torgerson, it was earlier suggested (cf Section 2.43) that the dimensions 'found' in MDS analyses should be regarded as virtual in nature, and dependent upon the judgement task and its perception by the individual. Such an approach would seem particularly appropriate in this case. Although the detailed organisation of all computer languages is strongly hierarchical, discussion concerning computer languages, e.g. an intra-language comparison, are not necessarily so. The current emphasis on the acquisition of 'top-down' programming knowledge might be seen to reflect the view of programming as problem perception rather than problem solving. The skill that the novice has to acquire is the ability to recognise particular 'patterns' in the specified problem. Most investigations in this knowledge domain have relied on some 'pictorial' differentiation between the hierarchical structures

derived from experts and novices. However using such a hierarchical knowledge representation, it is extremely difficult to differentiate between development of detailed expertise, and the possible development of a more abstract language of discourse about this detailed knowledge.

The success of the dimensional models of analysis in the identification of the development of such meta-dimensions is taken as a further indication of the importance of the consideration of synthetic reasoning in the individual's perception of similarity relations. These issues are discussed further in Section 5.5.

5.4 The perception of argument in text

"In retrospect, the distrust of some researchers regarding 'semantic' memory certainly seems justified. A more flexible and inclusive model of conceptual memory must deal with many more types of relations and with the effects of contexts of utilisation upon stored knowledge configurations ... I would submit that the study of textual processing might be a more productive means of gaining insights into knowledge and memory in realistic human conditions."

de Beaugrande (1980, p77)

5.4.1 Rationale

As was explained in Section 1.2, the original goal of this thesis was the intention to investigate and develop a methodology capable of mapping changes in 'personal' aspects of knowledge. Conventionally, studies of individual differences in perceived knowledge have taken 'snapshots' of concept relations within a static, structural representation of knowledge. The 'learning from text' studies in this field have primarily concerned

themselves with knowledge domains such as Newtonian physics and simple statistics, where the definition of concepts appears to be axiomatic, rather than part of the problem (e.g. Shavelson, 1972).

Recent work (e.g. Bisanz et al, 1979) has dealt with the more personal perception of story grammars, but only with regard to simple, concrete concepts or story character development. The concern here is to tap the effects of individuals 'making' meaning in terms of integrating various, possibly conflicting, sources of knowledge, which might include their own pre-existing knowledge and beliefs. Zajonc's (1955) ideas concerning the remodeling or 'tuning' of knowledge are relevant here, in that they involve reorganisation within an 'open' system. Some of the most interesting subject areas in these terms are the apparently 'hard' social-science topics such as economics and aspects of educational research. It is seemingly possible to construct value free theories and to deduce 'facts' in these fields, although in reality the **significance** of these facts is often entirely determined by implicit life goals. As for example in the topic of the texts chosen for this study, which concern the use of 'modern maths' in the infant and junior school curriculum.

Text manipulation studies concerned with the effects of argument structure or course development would obviously be facilitated by a methodology for repeatedly 'tapping' into the individual's perception of concept or topic relations. The requirements of the 'tapping' task are that it cover a wide area of knowledge, and yet take up as little time, and be as non-directive as possible. The various forms of similarity relation tasks have a considerable advantage over the conventional essay and multiple-choice form of questioning in these respects. A problem for studies attempting to tap changing patterns of concept relations following prose learning, is that the ideas and arguments in complex texts cannot be described simply in terms of the relations between concrete or theoretically primitive concepts. Elicitation techniques and analyses that were proven at the level of word structure cannot be **assumed** to be as effective in the examination of

complex prose structure. Theoretical developments have caused the 'nodes' in semantic networks to effectively enlarge. Current conceptions of semantic memory allow for whole frames or episodes as nodes. Although these may obviously be decomposed into their elements, we have argued in Chapter 3 that the natural form of processing of such entities is as whole units.

Given the desire to study learning from lengthy texts which have complex underlying structures, the researcher is confronted with the problem of constructing valid measures of understanding. Some of the problems involved in attempting to tap whether the reader has correctly perceived the author's general argument structure and the assigned relevance of different topics are discussed in Whalley (1982). It is extremely difficult to devise questions which measure how well the reader has integrated abstract conceptual issues, within the overly analytic (in the terms of our present argument) 'multiple choice' format. However it is also difficult to fit the writing and 'objective' marking of general essays within a formal experimental methodology. An intermediate task, which may solve both the practical and theoretical problems of experimental work with such material, is to require subjects to make similarity judgements concerning the important conceptual terms of the text.

For the domains of 'soft' knowledge that we have adopted, it is often not possible to attempt to identify, or to require readers to make judgements concerning, single concept nodes. Therefore in the study to be detailed, an attempt was made to cause readers to make relational judgements at the level of frames of experience; in this case in terms of short 'pen-pictures' describing a wide range of possible classroom activities. All the pen-pictures are contained in Appendix C, but some idea of their form may be gained from the first:

"Yasmin plays with a family of peg dolls she has made to go in the class Victorian dolls house and then writes up her diary about being a Victorian child."

The experiment deals with the readers' perception of argument in two texts, one making a case for, and one making a case against, the usefulness of 'modern maths'. The detailed results in terms of the readers' perceptions of text structure are not relevant here. However the success of the methodology as an experimental tool would further validate the general argument of the thesis, which emphasises the need to recognise the possibility of individuals adopting a holistic approach to judgements of similarity.

5.42 Method

Task The subjects were asked to read a four page text either in favour of or against the use of 'modern-maths' teaching in the junior school. These texts were respectively derived from the first chapter of a book by two well known advocates of learning maths through play activities (Holt and Dienes, 1973), and from a 'Black Paper' (Cox and Boyson, 1977) criticizing modern teaching methods. The texts spread over two A4 pages in the format used in the experiment, but are shown in a more condensed form in Appendix C. The subjects were told that they would receive a 'test' after reading the texts, and were asked to "underline the five sentences that you consider the most important in terms of the author's argument" whilst reading.

Before and after they had read the texts, the subjects were required to make similarity judgements between the twenty-two pen-pictures describing infant and junior classroom activities. It was intended that the pen-pictures should be thought of, and related to each other, as complete 'frames' of experience. In order to discourage any piecemeal decomposition of the pen-pictures, the subjects were asked to spend five minutes before the first reference-ranking session "visualising the scenes as they might occur in the classroom". They were also asked to continue thinking of the activities as "complete scenes" whilst making their relational judgements. The reference ranking similarity elicitation task was used, as detailed in the studies of Chapter 4, each subject being required to complete the task twice during the

course of the experiment. An incomplete data design was adopted because of time constraints. From a subset of eight items, each subject was randomly assigned four for each 'tapping' to serve as reference items. Appendix C contains the 22 pen-pictures used in the reference ranking task, and those used as reference items are indicated.

Design The intention of the study was to demonstrate the usefulness of this methodology for 'tapping' into changing patterns of conceptual relations. There were two parts to the study with two groups involved in each. In the first part, subjects were required to complete the reference ranking task before and after reading the text. One group received the pro-discovery method text (the pro-group), and the other the anti-discovery method text (the anti-group). In the second part, two different groups of subjects received in succession: the pro- or the anti-text; the reference ranking task; the anti- or the pro-text respectively; and then another reference ranking task. (These will be termed the 'pro-anti' and 'anti-pro' groups respectively.) Subjects attending each session were randomly assigned to groups. There were five subjects in all groups except the anti- which had four (due to the illness of one subject assigned to this group). Both parts of the study are of a 'mixed' design. The 'before and after' tapping involves 'within-subjects' aspects and the comparison between groups receiving different texts, or the same texts in a different order, involves 'between-subjects' aspects. Two data sets were obtained from each part of the study. In the first these correspond to a 'tapping' before and after reading a particular text, whilst in the second they correspond to a 'tapping' after the first text has been read and then again after the second.

Subjects and sessions The subjects were students at City of London Polytechnic taking a variety of courses; none were taking part in any educational studies courses. Subjects who signed up as being available were randomly assigned to either of two sessions. The first session, with the pro- and the anti- groups, took about forty minutes; for which the subjects

were paid £ 1.50. The second session, with the pro-anti and the anti-pro groups, took about one hour; for which the subjects were paid £ 2.

5.43 Results

A combined ALSCAL analysis was made of the four data sets. Therefore the relative positions and movements between each elicitation for each group are directly comparable, in that they have been subjected to the same normalization procedure before plotting. An 'individual differences' analysis was made with the configuration 'fixed' on the configuration obtained from an earlier study. This permitted stable analyses to be made with relatively small amounts of data, and also allowed comparisons to be made with a 'known' dimensional structure. The pilot study had involved a group of student teachers, who were broadly in favour of the 'discovery' method; making paired similarity judgements concerning all items in the set of pen-pictures. Within this configuration, pen-pictures describing 'discovery' maths activities were relatively highly weighted in terms of the maths dimension. In this analysis then, a high or low dimensional weighting on the 'maths' dimension may be interpreted as signifying more or less agreement respectively with the idea that the 'discovery' maths pen-pictures entail important mathematical ideas.

The previous trials using the pen-pictures indicated that subjects usually perceived three dimensions of variation in them: maths versus play; reading and writing versus activity based; and solitary versus group activity. Only the first two were relevant to the argument in the texts being studied in this case, and no apparent differences were found on the third. The figures (i.e. Figures 5.8 to 5.11) shown here then only represent the first two dimensions. It should be noted that the axis of the figures indicating the dimensional weightings have been foreshortened. The figures effectively represent the critical portion of what is normally an A3 sized plot. The thin lines represent the orientation of each subject's first elicitation

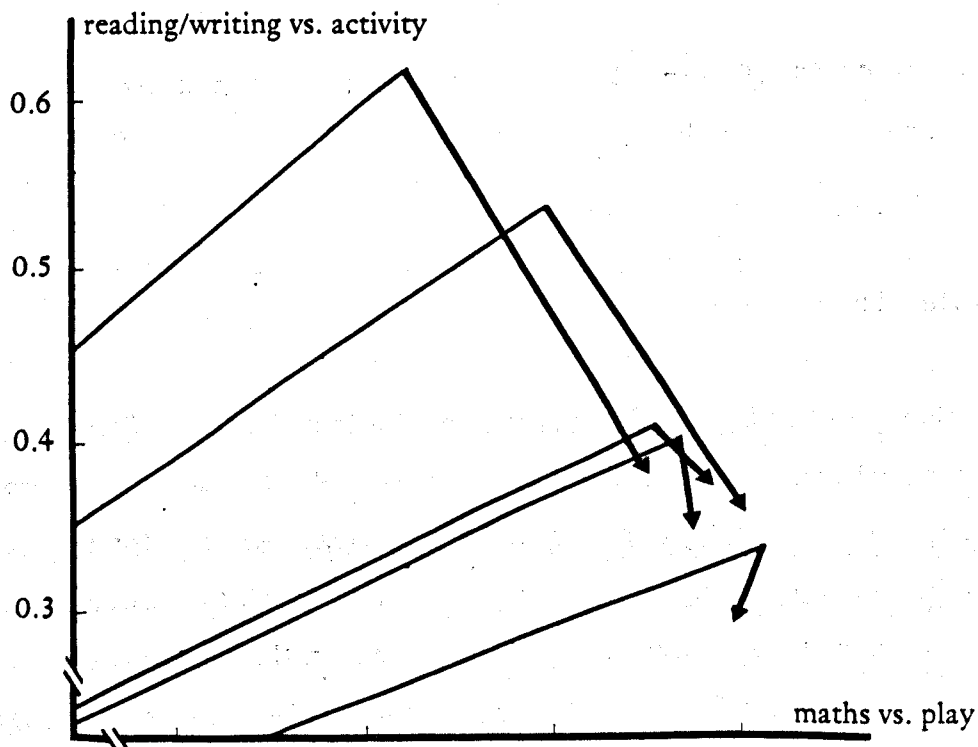


Figure 5.8 Dimensional weights for the group reading the pro-discovery method text (ALSCAL analysis, SSTRESS = 0.119, 22 stimuli).

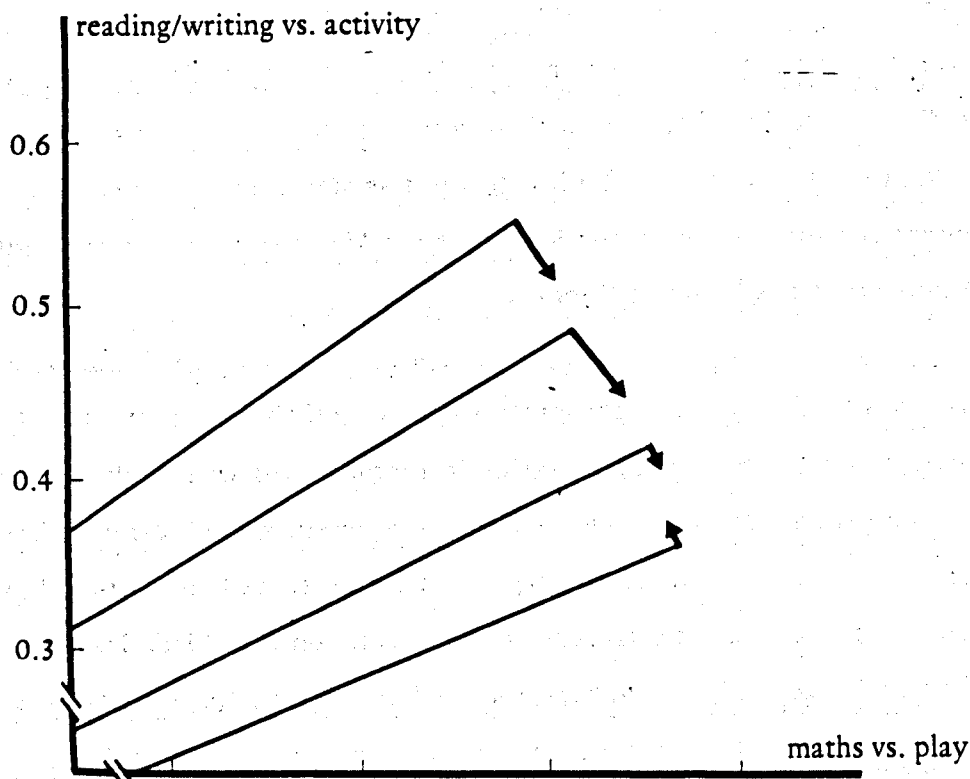


Figure 5.9 Dimensional weights for the group reading the anti-discovery method text (ALSCAL analysis, SSTRESS = 0.119, 22 stimuli).

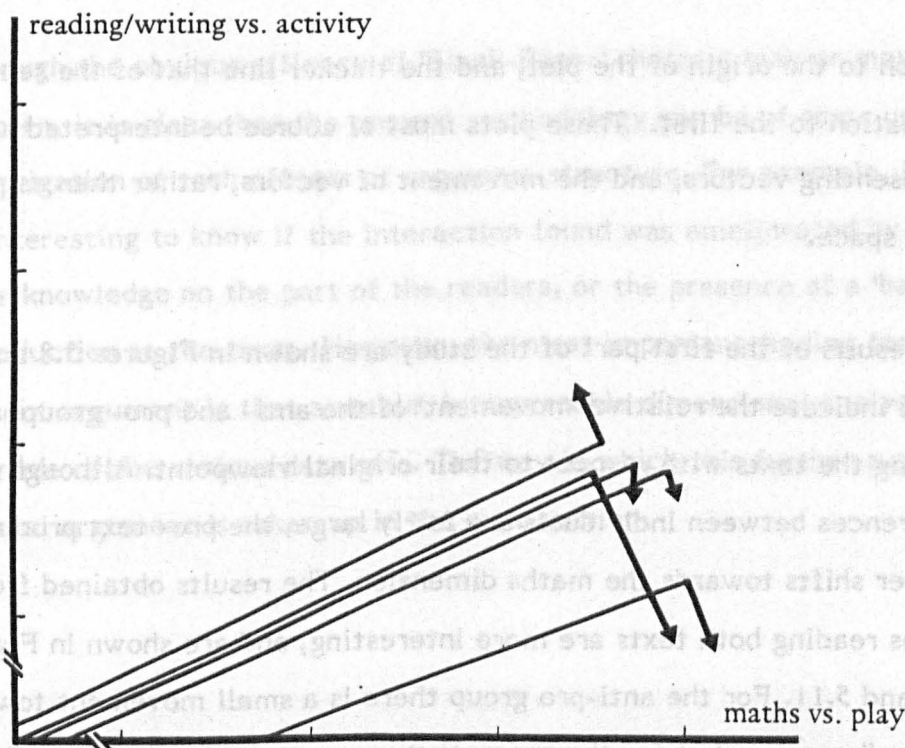


Figure 5.10 Dimensional weights for the group reading the anti-discovery method text followed by the pro-discovery method text (ALSCAL analysis, SSTRESS = 0.119, 22 stimuli).

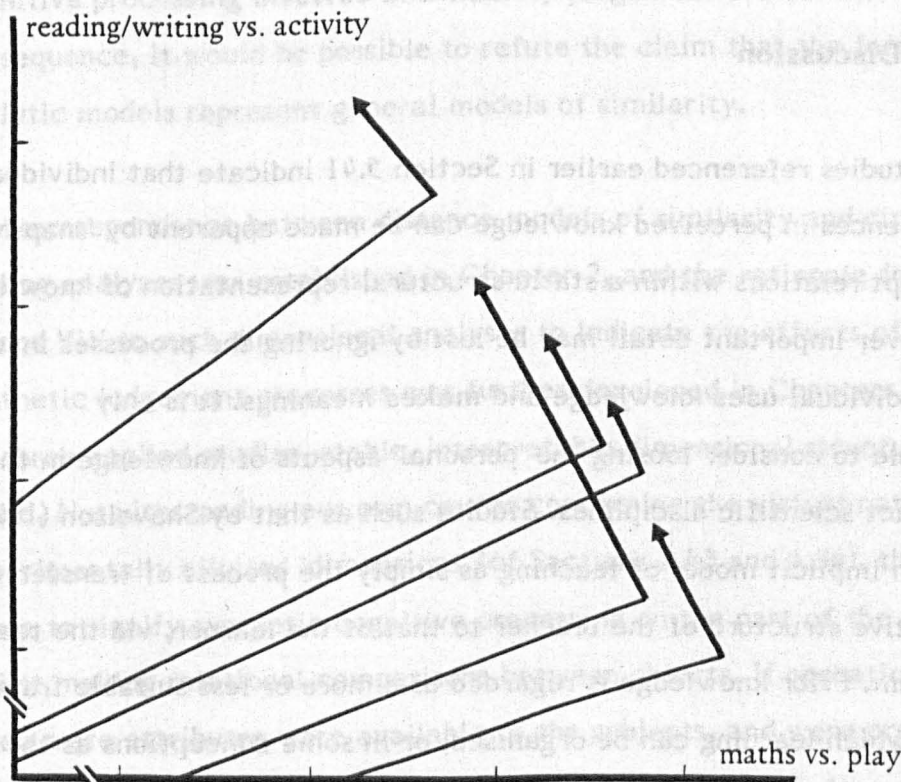


Figure 5.11 Dimensional weights for the group reading the pro-discovery method text followed by the anti-discovery method text (ALSCAL analysis, SSTRESS = 0.119, 22 stimuli)

session to the origin of the plot, and the thicker line that of the second elicitation to the first. These plots must of course be interpreted as representing vectors, and the movement of vectors, rather than as points in some space.

The results of the first part of the study are shown in Figures 5.8 and 5.9. These indicate the relative 'movement' of the anti- and pro- groups after reading the texts with respect to their original viewpoint. Although the differences between individuals are fairly large, the pro- text produces the greater shifts towards the maths dimension. The results obtained from the groups reading both texts are more interesting, and are shown in Figures 5.10 and 5.11. For the anti-pro group there is a small movement towards the maths dimension, but for the pro-anti there are very large shifts away from the maths dimension. It would seem that there is an interaction effect, in which the arguments against the discovery method are enhanced by a prior exposition in their favour.

5.44 Discussion

The studies referenced earlier in Section 5.41 indicate that individual differences in perceived knowledge can be made apparent by 'snapshots' of concept relations within a static structural representation of knowledge. However important detail may be lost by ignoring the processes involved when the individual uses knowledge and makes meanings. It is only possible to consider losing the 'personal' aspects of knowledge in the more abstract scientific disciplines. Studies such as that by Shavelson (1972) had an implicit model of teaching as simply the process of transferring the cognitive structure of the teacher to that of the learner, via the teaching medium. Prior knowledge is regarded as a more or less suitable framework upon which learning can be organised, or in some conceptions as the model of organisation likely to be imposed on new information. This naive 'instructional' model is clearly an inadequate description of learning from text, especially beyond the secondary level.

Although the obvious efficacy of 'Black-Paper' rhetoric may or may not be of concern, it is clear that the present methodology can be of some use in the investigation of such effects of argument structure. For example, it would be interesting to know if the interaction found was ameliorated by greater prior knowledge on the part of the readers, or the presence of a 'balanced' introduction to the texts. However, the most important finding for our present argument is that a stable, interpretable dimensional analysis could be made of the various data sets. The way in which this further supports our general argument is taken up in the next section.

5.5 General discussion

These three applied studies were made in order to further indicate how the purely analytic models are unable to adequately describe all forms of cognitive processing involved in similarity judgement tasks. As a consequence, it would be possible to refute the claim that the feature analytic models represent general models of similarity.

The correspondence between distance models of similarity and dimensional scaling analyses was established in Chapter 2, and the rationale for taking a good 'fit' to such dimensional analyses to indicate the effects of synthetic judgement processes was further developed in Chapters 3 and 4. In all three applied studies, stable, interpretable dimensional structures were found. Notwithstanding our own caveat concerning the virtual nature of experimentally elicited 'dimensions' (cf Sections 2.43 and 5.34), this is taken to signify synthetic cognitive processing on the part of the subjects, whilst making relational comparisons between objects. If operations in terms of discrete attributes were available to the subjects, and were preferred by them, then these results would not have been obtained. The feature analytic models of similarity judgement cannot account for this data.

These findings thus offer further proof of the lack of generality of the feature models of similarity, and lend support to our contention concerning the necessity to incorporate some aspect of synthetic cognitive process in any general model of similarity.

CHAPTER 6 Conclusion

6.1 A summary

"The capacity to 'think' in a rewarding sense seems to require not a little skill in keeping data, and several lines of thought, actively in mind long enough for their mutual implications to be developed associatively, logically, or both. This capacity, characteristic of contemplation or meditation in its literary and religious context, has tended to be neglected by psychologists." Reeves (1969, p20)

The underlying argument of this thesis, is that psychologists have chosen to examine a fundamentally important concept, that of similarity, with a distorting perspective. The underlying conceptual issues are not dealt with adequately, and only those experiments are made which fit the overly simplified representational assumptions of the dominant models. In a review of the psychological literature, it was noted that many of the early researchers appeared to be more aware of the epistemological problems confronting the reductionist approach to similarity, than is shown by the authors of the present day 'general theories' of similarity.

The fundamental idea involved in the model developed in Chapter 3 is that the global properties of referents are synthetically evaluated in terms of their contextual relations, whilst an analytic 'pattern matching' of local properties is made. Our discussion then centres round the argument that a truly general model of similarity must incorporate aspects of both analysis and synthesis. The feature models are shown to be based on an over simple representational model, and to consequently exclude from formal consideration all but the analytic cognitive operations. The experimental studies of Chapter 4 were designed to further highlight the inadequacies of the feature models of similarity. The empirical findings clearly indicate that global context effects, the perceived complexity of objects, and the interpretations made by subjects of their task cannot be ignored. The

applied studies of Chapter 5 were intended to show that synthetic relational judgements may in some cases be more appropriate than analytic forms of judgement. Stable, interpretable dimensional structures were found in these experiments, indicating the ability of the subjects to conceive of the particular stimuli in terms of their relational aspects. The corresponding analytic forms of judgement were shown to be less satisfactory.

An attempt was made in the early chapters to accommodate as wide a range of views from the relevant philosophical and psychological literatures as possible. However there must remain many other areas of psychology which have not been covered, and yet contain reasoned argument concerning the role and nature of similarity in human cognition. An example is Freud's discussion of the 'condensation' process of dream-thoughts. Here the distinction between similarity, consonance, approximation and the possession of common attributes is considered; processes which clearly correspond to those discussed here. Equally relevant are the problems that AI researchers have had to confront when trying to extend simple, feature checking, 'match' procedures to the immensely more complex comparisons required between knowledge and context sensitive schemas.

It is not possible here to remedy the relative inattention paid by psychological research to the important cognitive processes of contemplation and meditation. However in the next section an attempt will be made to encompass some of the implications of the acceptance of a notion of similarity that encompasses aspects of both analysis and synthesis.

6.2 Prospects and implications

"The verdict is not so suddenly arrived at, the proceedings only gradually merge into the verdict."

Franz Kafka (The Trial, 1970, p232)

It was noted in the development of the functional model, that a fundamental indeterminacy is apparently necessary in any knowledge representation system. The same aspect of an object may be treated as either 'attribute' or 'relation' depending on whether the object is the subject of analytic or synthetic cognitive operations. The objects chosen as the source of the experimental stimuli were necessarily rather simple, in order both that the link between behavioural protocols and elicited similarity judgements could more readily be argued, and also that the experiments parallel those found in the literature. A consequent shortcoming of the empirical investigations, is that they do not fully test all the representational and procedural implications of the model developed in Chapter 3. It is possible to envisage an appropriate development of this study in the richer domain of the relation between judgements of similarity and the comprehension of metaphor. Given that it could be argued that the distinction between analytic and synthetic cognition corresponds to the distinction between the comprehension of simile and metaphor, then experiments could be made with more complex linguistic stimuli. Similarity judgements made concerning, for example, a set of common household items should be differentially effected by their simultaneous use in some metaphoric context.

Several issues were noted in the review section that are important for the further development and testing of general models of similarity, although all are presently at an early stage of development in the literature. The transformational models of perceptual processes such as that proposed by Imai (1977) would be a useful component of any general model of similarity that attempted to take into account the global changing nature of relations. Takane (1981, and cf Section 2.43) has detailed the parallelism between the

initial data conversion process in the analysis models and the individual's representation system. The atheoretical treatment conventionally accorded to intransitive judgements was noted in Section 1.3., and an ideal analysis of similarities data would permit the investigator to model and test theoretical predictions concerning the way that such errors of judgement might arise. The attempt by Takane to incorporate a model of possible judgement processes within the data analysis procedure is a significant step in this direction.

The correspondence between synthetic and analytic modes of cognition and the developing conceptions of 'cognitive architecture' were noted in Chapter 3. The distinction made by Anderson (1983) between automatic and planned operations would clearly be important for models such as the one proposed here which are described in terms of the representation of objects and the possible operations on such representations. A development of Torgerson's ideas concerning the treatment of the dimensions resulting from MDS analysis¹⁵ was also considered in Sections 2.43 and 5.34. It was suggested that the 'dimensions' may most usefully be regarded as pure artifacts of the task, and the individual's conception of it. A consequence of adopting this approach, is that it allows the use of this analytic technique to escape the dominating influence of its long psychophysical tradition. Its use may then be extended to examine a wider sphere of cognitive activity.

Just such a use is detailed in the last study of Chapter 5 where an analysis was made of successive mappings of subjects' concept relations. It was noted in Section 1.2 that the original goal of this thesis was the wish to validate and empirically test a methodology with which investigations of the growth of human knowledge might be made. Within the field of linguistics, Saussure (1960) developed the important distinction between 'synchronic' and 'diachronic' language studies. Synchronic studies would correspond to the attempt to tap an individual's pattern of concept relations at one instant in time. Proponents of the diachronic methods of study would argue that this must be a conceptually inadequate description of the individual's reality,

and that language, and thus knowledge, can only be examined within the context of its growth and development.

In Chapters 2 and 3, some of the criticisms made of cognitive psychology by philosophers such as Putnam (1981) and Fodor (1983) were detailed.

Essentially they are claiming that cognitive psychology has adopted an inappropriate model of the mind; the 'mind as a computer' metaphor. The arbitrary level of symbolism from which the models of AI and cognitive psychology start is taken to indicate that they are really only models of syntax rather than semantics. In the light of this criticism, the attempt to incorporate synthetic judgements within the general scheme of cognition can be seen to follow from a view of the mind as an 'interpreting engine' rather than as an 'inferencing engine'. The scope of the argument was consequently broadened in Chapter 3 to encompass the distinction, made in relation to metaphor comprehension, between the 'constructivist' and 'non-constructivist' perceptions of language. The correspondence between the proposed model of similarity and Black's (1962) functional analysis of how metaphors 'work' being noted.

The studies detailed here were intended to show that the scaling methodologies, when based on an appropriately general model of similarity, may contribute to other than a static synchronic description of human knowledge. It is hoped that all the developments outlined might be brought together to further support William James' original contention that psychology, or in the case of our limited argument, models of similarity, "- must be writ both in synthetic and in analytic terms."

References

- Anderson, J.R. 1976 **Language, memory and thought**, Erlbaum, Hillsdale, New Jersey.
- Anderson, J.R. 1983 **The Architecture of Cognition**. Harvard University Press, Cambridge, Mass.
- Anderson, J.R. and Bower, G.H. 1973 **Human associative memory**. Wiley: London.
- Arnold, J.B. 1971 **A multidimensional scaling study of semantic distance**. *Journal of Experimental Psychology Monograph*, 90 (2), 349-372.
- Attneave, F. 1950 **Dimensions of similarity**. *American Journal of Psychology*, 63 (4), 516-556.
- Bartlett, F.C. 1932 **Remembering - a study in experimental and social psychology**. Cambridge University Press.
- Beals, R., Krantz, D. and Tversky, A. 1968 **Foundations of multidimensional scaling**. *Psychological Review*, 75 (2), 127-142.
- Beaugrande, R. de 1980 **Text, discourse, and process: toward a multidisciplinary science of texts**. London: Longman.
- Bertalanffy, L. von 1973 **General System Theory**. Penguin Books.
- Bisanz, G.L., Laporte, R.E., Vesonder, G.T. and Voss, J.F. 1979 **On the representation of prose: new dimensions**. *Journal of Verbal Learning and Verbal Behavior*, 18 (2), 187-198.
- Black, M. 1962 **Metaphor**. In: Black, M. *Models and Metaphors*, Ithaca, N.Y.: Cornell University Press.
- Brentano, F. 1874 **Psychology from an empirical standpoint**. Leipzig, (Republished: Ed. L.L. McAlister, Routledge 1973).
- Brooks, L. 1978 **Nonanalytic concept formation and memory for instances**. In: Rosch, E. and Lloyd, B.B., *ibid*.
- Bush, R.R. and Mosteller, F. 1951 **A model for stimulus generalization and discrimination**. *Psychological Review*, 58, 413-423.

- Butler, A. (Bishop) 1910 **Works**. Edited by W.E. Gladstone, 2nd Vol., London.
- Carroll, J.D. and Chang, J.J. 1970 **Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition**. *Psychometrika*, 35, 283-319.
- Chernoff, H. and Haseeb, R.M. 1973 **The use of faces to represent points in K-dimensional space graphically**. *Journal of the American Statistical Association*, 68, 361-368.
- Collins, A.M. and Loftus, E.F. 1975 **A spreading-activation theory of semantic processing**. *Psychological Review*, 82, 407-428.
- Collins, A.M. and Quillian, M.R. 1969 **Retrieval time for semantic memory**. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-248.
- Coombs, C.H. 1964 **A theory of data**. New York: Wiley.
- Cox, C.B. and Boyson, R. (eds) 1977 **Black Paper 1977**. Temple Smith, London.
- Coxon, A.P.M. and Jones, C.L. 1979 **Measurement and Meaning**. London: Macmillan.
- Cunningham, J.P. and Shepard, R.N. 1974 **Monotone mapping of similarities into a general metric space**. *Journal of Mathematical Psychology*, 11, 335-363.
- Dasgupta, S. 1968 **A history of Indian philosophy**. Cambridge.
- de Leeuw, J. and Heiser, W.J. 1980 **Multidimensional scaling with restrictions on the configuration**. In: Krisnaiah, P.R. (ed), *Multivariate analysis*, Vol 5., Amsterdam, North Holland.
- Dearborn, G.V.N. 1910 **Notes on the discernment of likeness and of unlikeness**. *Psychological Review*. 7 (3), 57-64.
- Deese, J. 1962 **On the structure of associative meaning**. *Psychological Review*, 69 (3), 161-175.
- Deshpande, M. 1972 **On the notion of similarity in Indian poetics**. *Journal of Indian Philosophy*, 2, 21-52.
- Dunlap, K. 1912 **The nature of perceived relations**. *Psychological Review*, XIX (6), 415-446.

- Ekman, G. 1954 **Dimensions of colour vision** *Journal of Psychology*, 38, 467-474.
- Ekman, G. and Sjöberg, L. 1965 **Scaling**. *Annual Review of Psychology*, 16, 451-474.
- Flugel, J.C. 1933 **A hundred years of psychology**. London: Duckworth.
- Fodor, J.A. 1983 **Modularity of mind**. MIT Press, Mass. and London.
- Galanter, E.H. 1956 **An axiomatic and experimental study of sensory order and measure**. *Psychological Review*, 63, 16-28.
- Gardner, W.R. 1978 **Aspects of a stimulus: features dimensions and configurations**. In: Rosch, E. and Lloyd, B.B., *ibid*.
- Gregson, R.A.M. 1975 **Psychometrics of similarity**. New York: Academic Press
- Gregson, R.A.M. 1976 **A comparative evaluation of seven similarity models**. *British Journal of Mathematical Statistical Psychology*, 29, 139-156.
- Guthrie, W.K.C. 1967 **A history of Greek philosophy**. Cambridge University Press.
- Hollan, J.D. 1975 **Features and semantic memory set-theoretic or network model?** *Psychological Review*, 82 (2), 154-155.
- Hollingworth, H.L. 1913 **Judgements of similarity and difference**. *Psychological Review*, 20, 271-289.
- Holman, E.W. 1972 **The relation between hierarchical and euclidean models for psychological distances**. *Psychometrika*, 37 (4), 417-424.
- Holt, M. and Dienes, Z. 1973 **Let's play maths**. Penguin
- Huskisson, E.C. 1974 **Measurement of pain**. *The Lancet*, 2, 1127-1131.
- Imai, S. 1977 **Pattern similarity and cognitive transformations**. *Acta Psychologica*, 41, 433-447.
- James, W. 1890 **The principles of psychology**. Henry Holt, (Republished: Dover, 1950).
- James, W. 1893 **Mr. Bradley on immediate resemblance**. *Mind*, 2, 208-210.

- Junge, K. 1978 **A general parametric model of similarity and subjective difference.** *Scandinavian Journal of Psychology*, 19, 111-119.
- Kafka, F. 1970 **The Trial.** Penguin Books.
- Katz, J.J. and Fodor, J.A. 1963 **The structure of a semantic theory.** *Language*, 39, 170-210.
- Kahney, H. 1982 **An in-depth study of the cognitive behaviour of novice programmers.** PhD Thesis, The Open University.
- Kelly, G.A. 1955 **The Psychology of Personal Constructs.** New York: Norton.
- Krantz, D.H. 1972 **A theory of magnitude, estimation and cross-modality matching.** *Journal of Mathematical Psychology*, 9, 168-199.
- Krantz, D.H. and Tversky, A. 1975 **Similarity of rectangles: an analysis of subjective dimensions.** *Journal of Mathematical Psychology*, 12, 4-34.
- Krumhansl, C.L. 1978 **Concerning the applicability of geometric models to similarity data: the interrelationship between similarity and spatial density** *Psychological Review*, 85 (5), 445-463.
- Kruskal, J.B. 1964 **Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis.** *Psychometrika*, 29 (1).
- Kruskal, J.B. and Wish, M. 1978 **Multidimensional Scaling.** Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-001. Beverly Hills and London: Sage Publications.
- Kulpe, O. 1895 **Outlines of psychology.** London: Swan Sonnenschein.
- Lakoff, G. and Johnson, M. 1980 **Metaphors we live by.** University of Chicago Press.
- Lingoes, J.C. 1979 **Measurement.** In: Lingoes, J.C. et al, *ibid.*
- Lingoes, J.C. and Borg, I. 1978 **A direct approach to individual differences scaling using increasingly complex transformations.** *Psychometrika*, 43 (4), 491-519.
- Lingoes, J.C., Roskam, E.E., and Borg, I. 1979 **Geometric representations of relational data.** Mathesis Press, Ann Arbor, Michigan.

- Luce, R.D., Bush, R.R. and Galanter, E. 1963 **Handbook of mathematical psychology**. Vol I, New York: Wiley.
- MacCullum, R. 1981 **Evaluating goodness of fit in nonmetric multidimensional scaling by ALSCAL**. *Applied Psychological Measurement*, 5 (3), 377-382.
- Mccabe, A.K. 1980 **A rhetoric of metaphor**. PhD Thesis, University of Virginia.
- McDougall, W. 1923 **An outline of psychology**. London: Methuen.
- Meddis R. 1975 **Statistical Handbook for non-statisticians**, McGraw-Hill.
- Melzack, R. 1975 **The McGill pain questionnaire: major properties and scoring methods**. *Pain*, 277-299.
- Metzler, J. and Shepard, R.N. 1974 **Transformational studies of the internal representation of three-dimensional objects**. In: R. Solso (ed.) 'Theories in cognitive psychology: the Loyola Symposium'. Potomac, Md.: Erlbaum.
- Miller, G.A. and Johnson-Laird, P.N. 1976 **Language and perception**. Cambridge, Harvard University Press.
- Minsky, M. 1975 **A framework for representing knowledge**. In: Winston, P. (ed), **The Psychology of computer vision**, New York: McGraw Hill.
- Nadin, M. 1980 **The logic of vagueness and the category of synecism**. *The Monist*, 63 (3), 351-363.
- Naveh-Benjamin, M. and Pachella, R.G. 1982 **The effect of complexity on interpreting "Chernoff" faces**. *Human Factors*, 24 (1), 11-18.
- Neal, D.E., Whalley, P.C., Flowers, M.W. and Wilson, D.H. 1981 **The effects of an adherent polyurethane film and conventional absorbent dressing in patients with small partial thickness burns**. *The British Journal of Clinical Practice*, 35 (7/8), 254-257.
- Neisser, V. 1976 **Cognition and reality**. San Francisco: Freeman.
- Noble, C.E. 1957 **Psychology and the logic of similarity**. *The Journal of General Psychology*, 57, 23-43.
- Ogden, C.K. 1967 **Opposition - A linguistic and psychological analysis**. Indiana University Press, Bloomington.

- Ortony, A. (ed) 1979 **Metaphor and thought**. Cambridge University Press.
- Ortony, A. 1979 **Metaphor: a multidisciplinary problem**. In: Ortony, A., *ibid*.
- Osgood, C.E., Suci, G.J. and Tannenbaum, P.H. 1957 **The measurement of meaning**. University of Illinois Press, Urbana, Chicago.
- Osherson, D.N. and Smith, E.E. 1982 **On the adequacy of prototype theory as a theory of concepts**. *Cognition*, 12 (3), 299-318.
- Preece, P.F.W. 1981 **Mapping cognitive structure: a comparison of methods**. *Journal of Educational Psychology*, 68 (1), 1-8.
- Putnam, H. 1981 **Reductionism and the nature of psychology**. In: Haugeland, J. (ed), *Mind design, philosophy, psychology*, A.I., Bradford Books, Publishers, Inc. Montgomery, VT
- Pylyshyn, Z.W. 1979 **Metaphorical imprecision and the "Top-Down" research strategy**. In: Ortony, A., *ibid*.
- Quine, W. 1953 **Two dogmas of empiricism**. In: *From a logical point of view*, Cambridge, Mass., Harvard University Press.
- Ramsay, J.D. 1977 **Maximum likelihood estimation in multidimensional scaling**. *Psychometrika*, 42 (2), 241-266.
- Reeves, J.W. 1969 **Thinking about thinking**. Methuen
- Restle, F. 1959 **A metric and an ordering on sets**. *Psychometrika*, 24 (3), 207-220.
- Richards, I.A. 1936 **The philosophy of rhetoric**. London: Oxford University Press.
- Richardson, M.W. 1938 **Multidimensional psychophysics**. *Psychological Bulletin*, 35, 659.
- Rips, L.J., Schoben, E.J. and Smith, E.E. 1973 **Semantic distance and the verification of semantic relations** *Journal of Verbal Learning and Verbal Behavior*, 12, 1-20.
- Rosch, E. 1975 **Cognitive reference points**. *Cognitive Psychology*, 7, 532-547.
- Rosch, E. 1978 **Principles of categorization**. In: Rosch, E. and Lloyd, B.B., *ibid*.

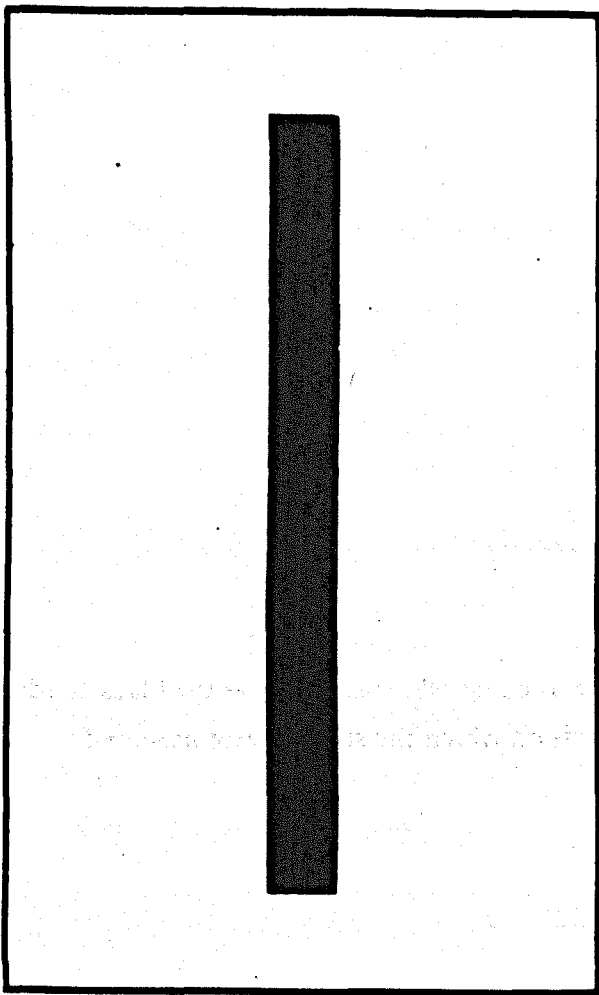
- Rosch, E. and Lloyd, B.B. (eds) 1978 **Cognition and categorization.** Erlbaum, Hillsdale, New Jersey.
- Rosen, S. 1980 **The limits of analysis.** New York: Basic Books.
- Roskam, E.E. 1979 **The nature of data: interpretation and representation.** In: Lingo, J.C. et al, *ibid.*
- Ross, W.D. (ed) 1928 **The works of Aristotle.** Oxford.
- Russell, B. 1903 **The principles of mathematics.** George Allen and Unwin.
- Russell, B. 1940 **An inquiry into meaning and truth.** George Allen and Unwin Ltd.
- Saussure, F. 1960 **Course in general linguistics.** London: Peter Owen
- Schiffman, S.S., Reynolds, M.L., and Young, F.W. 1981 **Introduction to multidimensional scaling.** New York: Academic Press.
- Schoenfeld, A.H. and Herman, D.J. 1982 **Problem perception and knowledge structure in expert and novice mathematical problem solvers.** *Journal of Experimental Psychology: Learning, Memory & Cognition*, 8 (5), 484-494.
- Schonemann, P.H. 1977 **Similarity of rectangles.** *Journal of Mathematical Psychology*, 16, 161-165.
- Schonemann, P.H., James, W.L. and Carter, F.S. 1979 **Statistical inference in Multidimensional Scaling: a method for fitting and testing Horan's Model.** In: Lingo, J.C. et al, *ibid.*
- Schroeder, Driver and Streufert. 1967 **Human Information Processing.** Holt Rinehart and Winston.
- Shavelson, R.J. 1972 **Some aspects of the correspondence between content structure and cognitive structure in physics instruction.** *Journal of Educational Psychology*, 63 (3), 225-234.
- Shepard, R.N. 1962 **The analysis of proximities: multidimensional scaling with an unknown distance function.** *Psychometrika*, 27 (2).
- Shepard, R.N. 1980 **Multidimensional scaling, tree fitting and clustering.** *Science*, 210, 390-398.

- Shepard, R.N. and Arabie, P. 1979 **Additive clustering: representation of similarities as combinations of discrete overlapping properties.** Psychological Review, 86 (2), 87-123.
- Shepard, R.N., Kilpatrick, D.W. and Cunningham, J.P. 1975 **The internal representation of numbers.** Cognitive Psychology, 7, 82-138.
- Smith, E.E., Shoben, E.J. and Rips, L.J. 1974 **Structure and process in semantic memory: a featural model for semantic decisions.** Psychological Review, 81 (3), 214-241.
- Spencer, H. 1881 **The principles of psychology.** Vol II, 3rd Edn., Williams and Norgate, London.
- Takane, Y. 1981 **Multidimensional successive categories scaling: a maximum likelihood method.** Psychometrika, 46 (1), 9-28.
- Takane, Y. and Carroll, J.D. 1981 **Nonmetric maximum likelihood multidimensional scaling from directional rankings of similarities.** Psychometrika, 46 (4), 389-405.
- Takane, Y., Young, F.W. and de Leeuw, J. 1977 **Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features.** Psychometrika, 42, 7-67.
- Torgerson, W.S. 1958 **Theory and methods of scaling.** New York: Wiley.
- Torgerson, W.S. 1965 **Multidimensional scaling of similarity.** Psychometrika, 30 (4), 379-393.
- Torgerson, W.S. 1983 **The ideal type model.** Unpublished MS, The John Hopkins University.
- Tucker, L.R. 1960 **Intra-individual and inter-individual multidimensionality.** In: **Psychological scaling: theory and applications**, Gulliksen, H. and Messick, S. (eds) New York: Wiley
- Tversky, A. 1977 **Features of similarity.** Psychological Review, 84 (4), 327-352.
- Tversky, A. 1978 **Studies of similarity.** In: Rosch, E. and Lloyd, B.B., *ibid.*
- Tversky, A. and Gati, I. 1982 **Similarity, separability and the triangle inequality.** Psychological Review, 89 (2), 123-154.

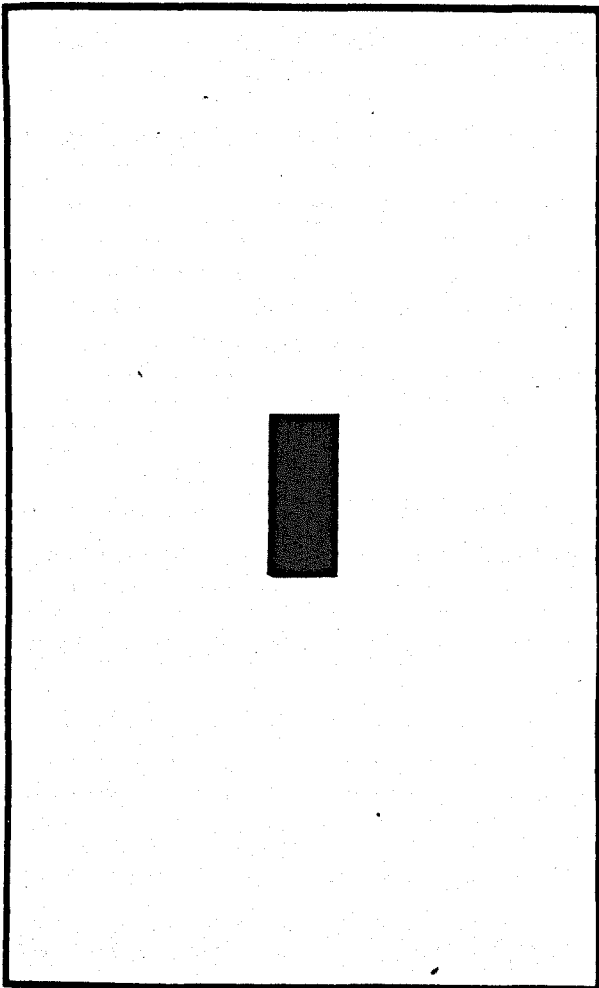
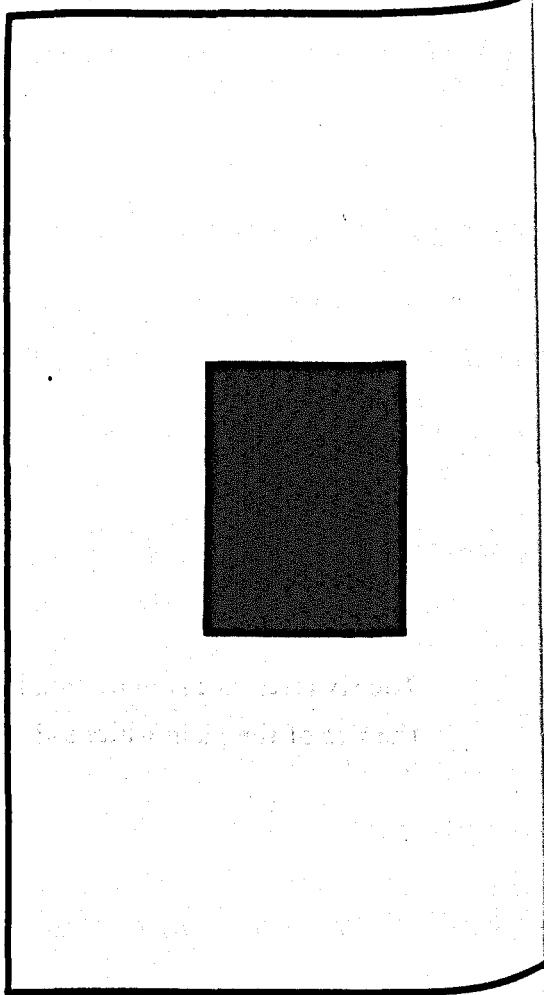
- Tversky, A. and Krantz, D.H. 1969 **Similarity of schematic faces: a test of interdimensional additivity.** *Perception and Psychophysics*, 5, 124-128.
- Wainer, H. and Kaye, K. 1974 **Multidimensional scaling of concept learning in an introductory course.** *Journal of Educational Psychology*, 66 (4), 591-598.
- Wallach, M.A. 1958 **On psychological similarity.** *Psychological Review*, 65 (2), 103-116.
- Whalley, P.C. 1982 **Argument in text and the reading process** In: Flammer, A. and Kintsch, W. (eds), **Discourse Processing**, North Holland.
- Winograd, T. 1975 **Frame representations.** In: Bobrow, D.G. and Collins, A. (eds), **Representation and Understanding.**
- Wittgenstein, L. 1968 **Philosophical investigations.** Blackwell.
- Young, F.W. 1975 **Scaling replicated conditional rank - order data.** *Sociological Methodology*, 129-170.
- Zajonc, R.B. 1955 **Cognitive structure and cognitive tuning.** PhD Thesis, University of Michigan.
- Zinnes, J.L. and MacKay, D.B. 1983 **Probabilistic multidimensional scaling: complete and incomplete data.** *Psychometrika*, 48 (1), 27-48.

Appendix A

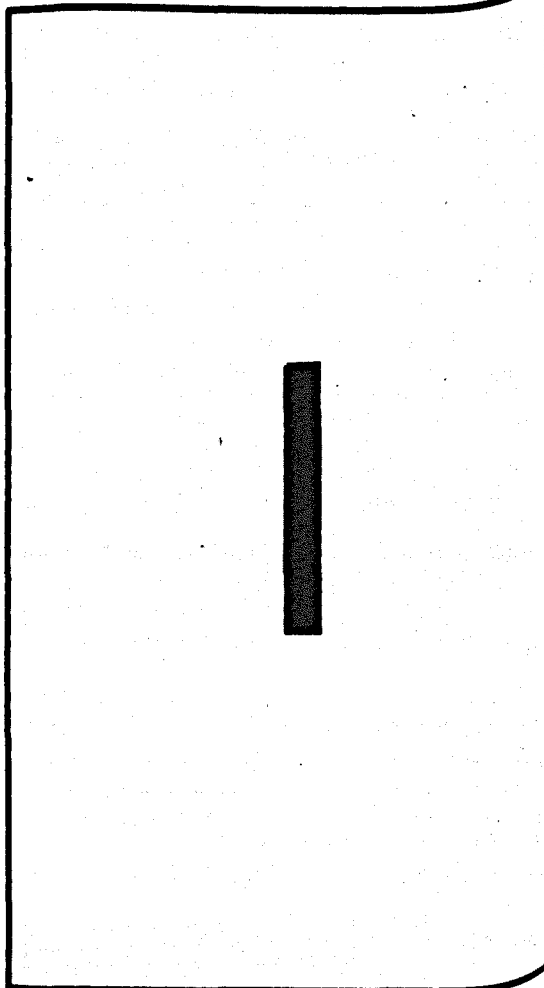
The six reference stimuli used in the 'rectangles' experiments — the black borders indicate the size of the plain white index cards on which the stimuli were mounted.

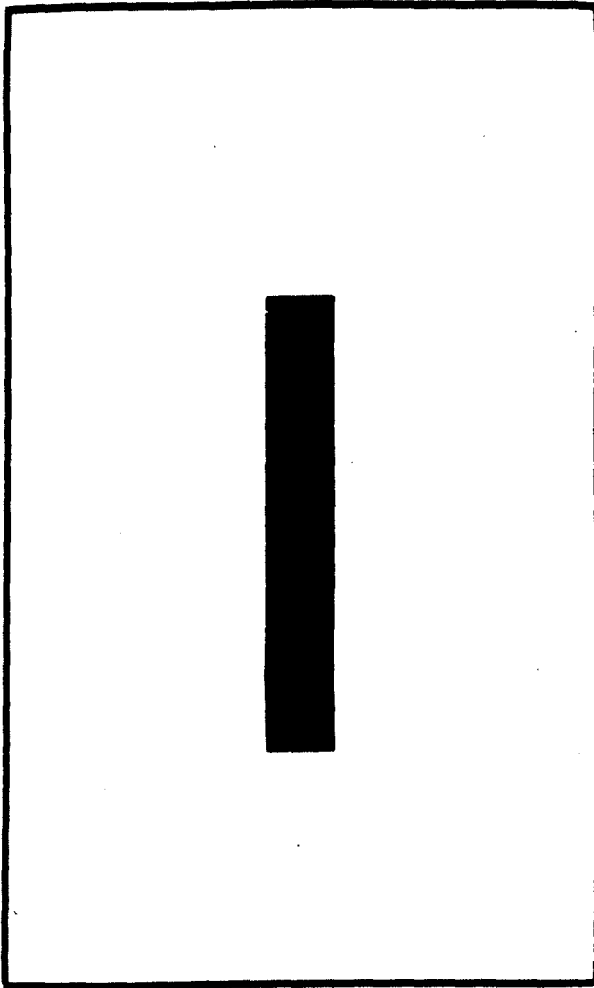


5

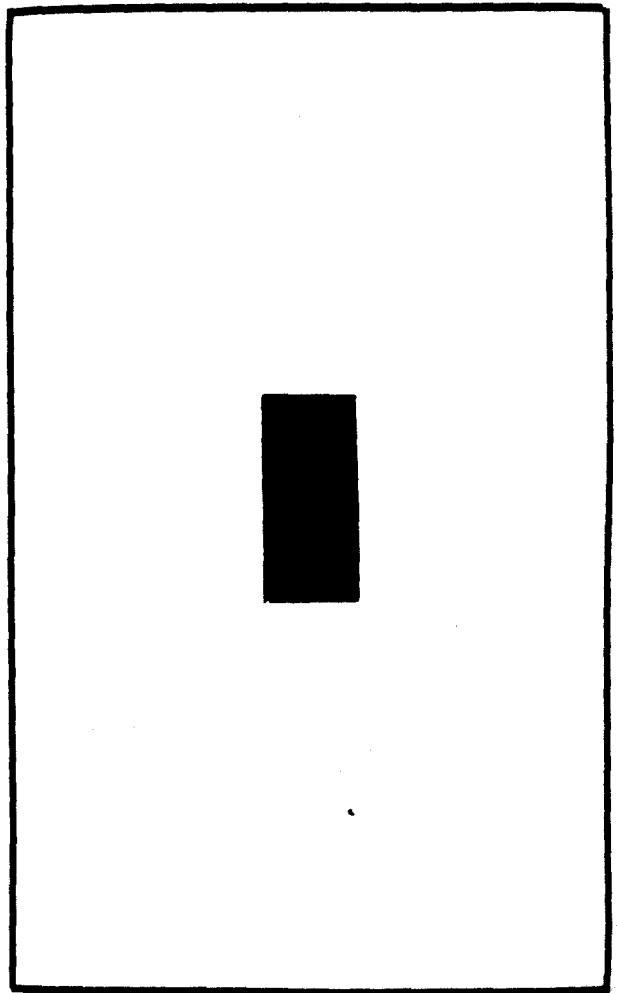


9





13



16

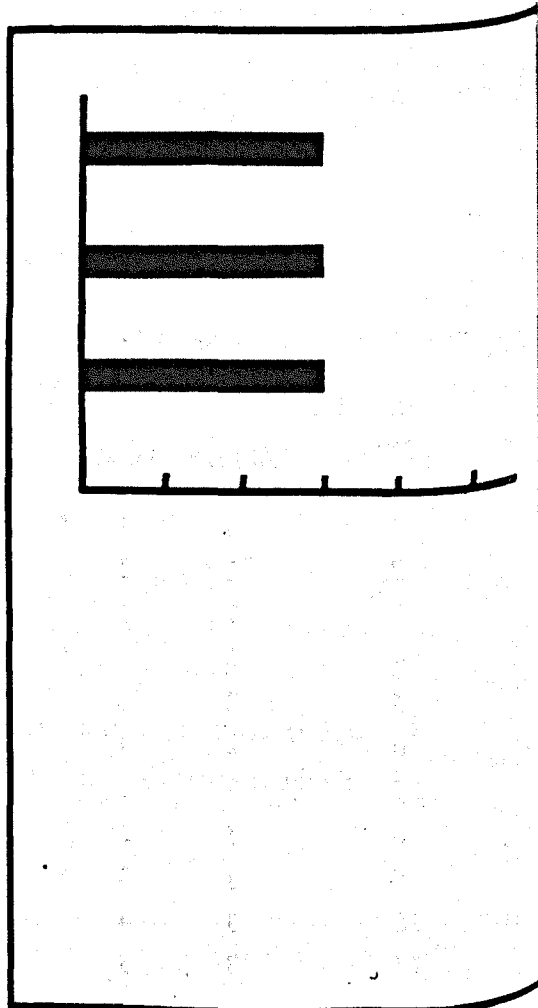
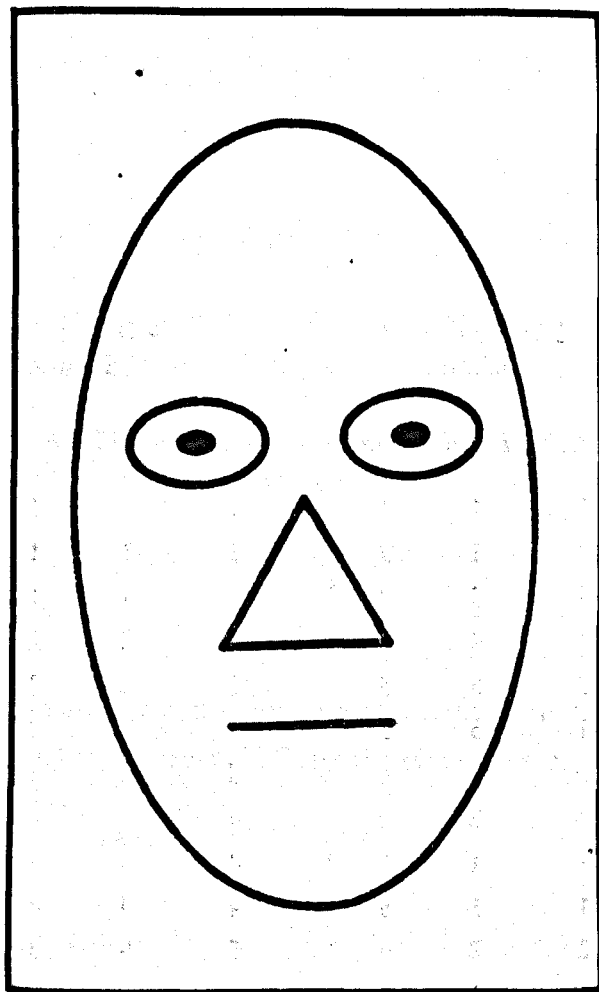
Appendix B

Bar chart and cartoon face representations of reference stimuli numbers 3, 6, 9 and 11 - the black borders indicate the size of the plain white index cards on which the stimuli were mounted.

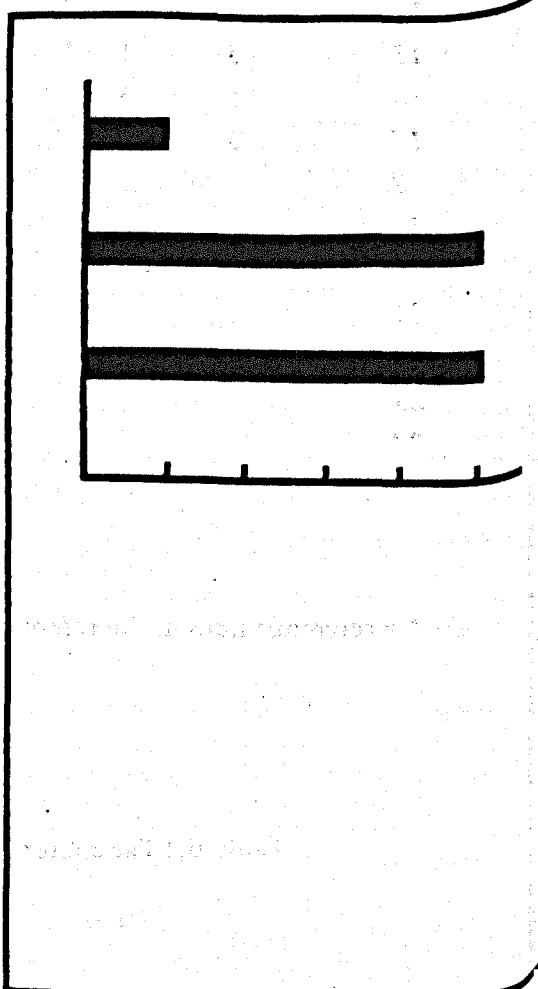
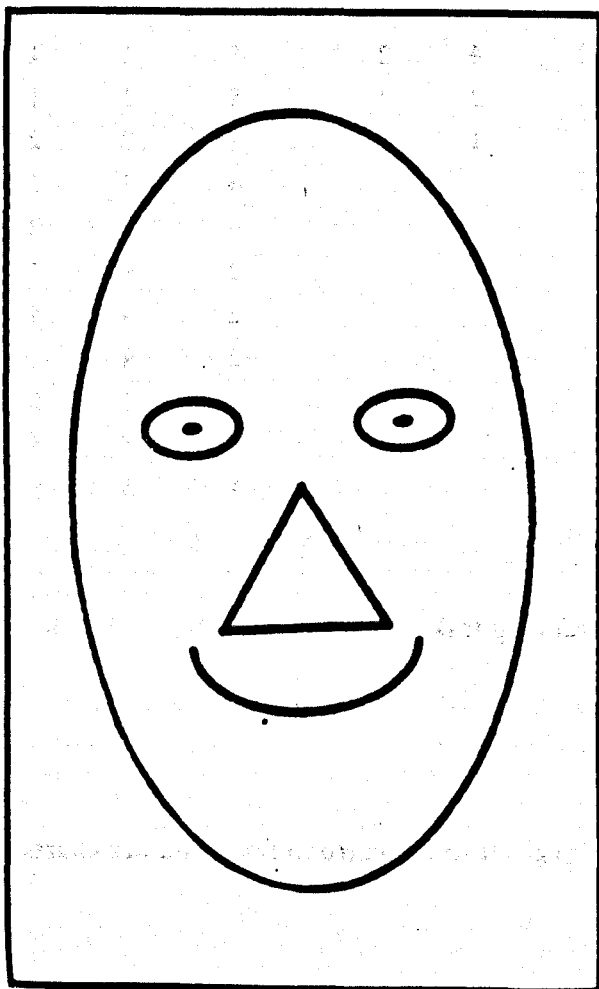
Stimulus No.	2 Dimensional Stimuli			3 Dimensional Stimuli			4 Dimensional Stimuli			
	<i>Mouth</i>	<i>Pupil</i>		<i>Mouth</i>	<i>Pupil</i>	<i>Eye</i>	<i>Mouth</i>	<i>Pupil</i>	<i>Eye</i>	<i>Nose</i>
1	1	1	*	1	1	1	1	1	1	3
2	1	3		2	1	2	1	1	1	1
3	1	5		3	3	3	1	1	1	5
4	2	2	*	4	4	4	2	2	2	3
5	2	3		5	5	5	2	2	2	2
6	2	4		1	5	5	2	2	2	4
7	3	1		2	4	4	3	3	3	3
8	3	2		4	2	2	4	4	4	3
9	3	3	*	5	1	1	4	4	4	2
10	3	4		1	1	5	4	4	4	4
11	3	5		2	2	4	5	5	5	3
12	4	2		4	4	2	5	5	5	1
13	4	3		5	5	1	5	5	5	5
14	4	4	*	1	5	1	5	1	1	3
15	5	1	*	2	4	2	5	1	1	1
16	5	3		4	2	4	5	1	1	5
17	5	5		5	1	5	4	2	2	3
18							4	2	2	2
19							4	2	2	4
20							2	4	4	3
21							2	4	4	2
22							2	4	4	4
23							1	5	5	3
24							1	5	5	1
25							1	5	5	5

* = reference item in the reference ranking task

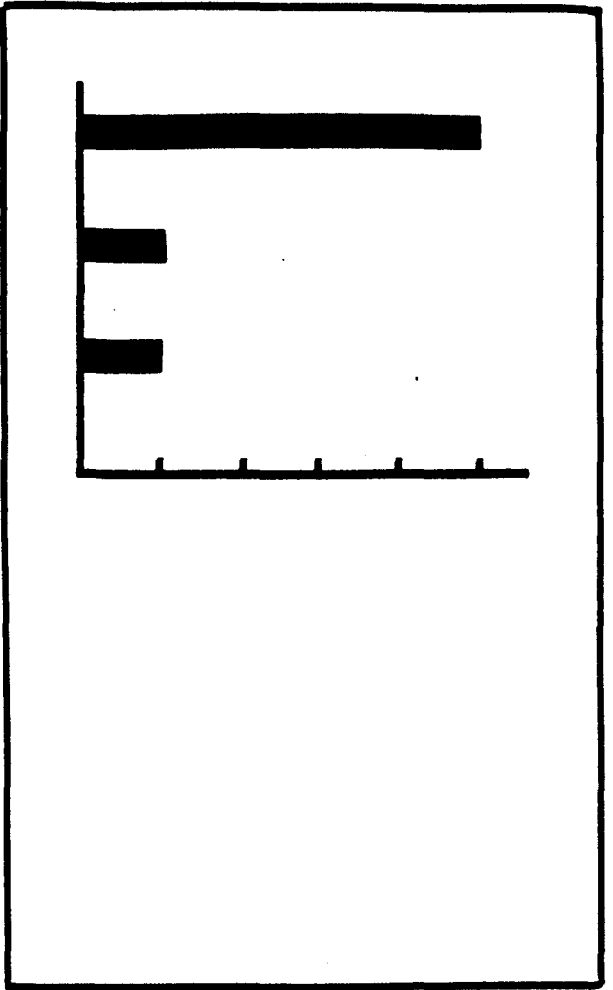
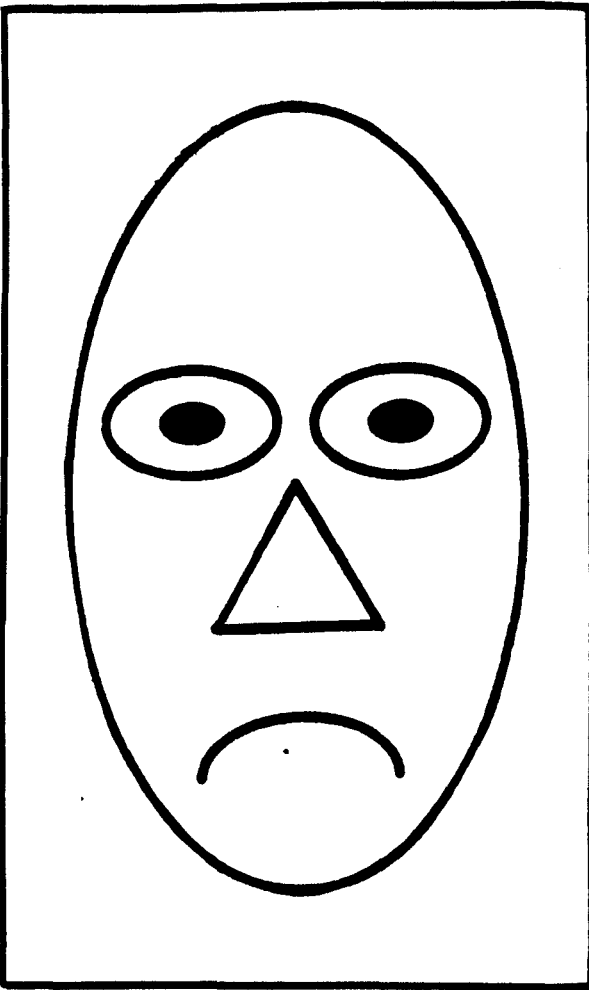
Table B.1 Parameter values given to the cartoon faces and bar charts



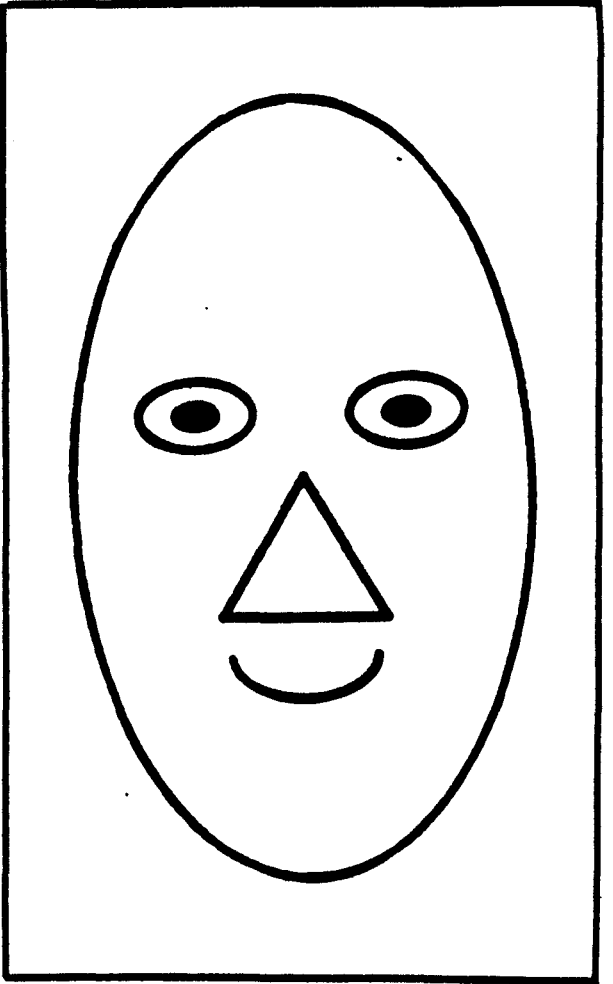
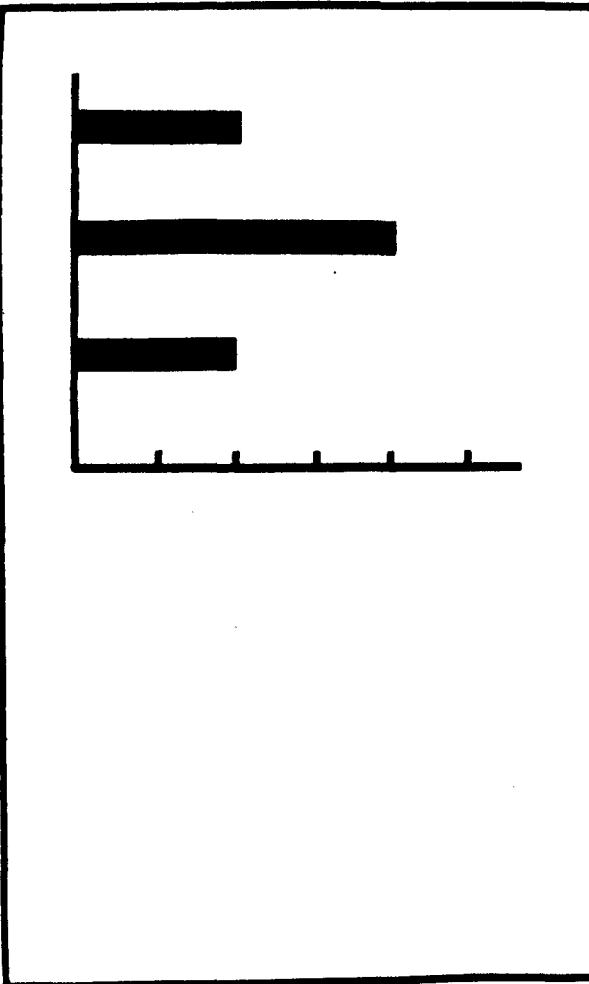
The two representations of stimulus number 3



The two representations of stimulus number 6



The two representations of stimulus number 9



The two representations of stimulus number 11

Appendix C

The 22 'pen pictures' describing scenes in a junior school classroom.

1. Yasmin plays with a family of peg dolls she has made to go in the class Victorian dolls house and then writes up her diary about being a Victorian child.
2. Diane and Jane go round the class asking each child what colour toothbrush they use. They record the results on a graph and at the end of the day they show the graph to the rest of the class and explain what they have found out.
3. John is making a calendar and flicks through the months saying each one to himself, "January, February, March....".
4. Nicola has listened to her teacher telling a story about a hamster and she is now writing her own hamster escape story.
5. Whilst Alison is painting a picture Jane flicks paint at her. The blue merges into the red, and Alison changes her picture into a rainbow where lots of colours blend together.
6. Sahara and Nigel are cutting out equilateral triangles from sticky paper and fitting them together to make different patterns. Then they write about what they have done.
7. The class rabbit was allowed to roam freely around the room for fifteen minutes and the children were encouraged to observe her movements and activities. They talked to each other and exchanged ideas while this was going on, and were then asked to write as accurate a description as possible of her.
8. Simon is making a balsa wood model of a boat that Robinson Crusoe could have used.
9. Angela and Betty write out the seven times and eight times tables and check their answers together.
10. Joane is painting a picture of her house chattering to herself as she goes along about where her room is and where the stairs go, etc.
11. Serena is trying to work out the surface of the class guinea pig by cutting out a paper cylinder that fits closely round it.
12. Claire and Angela are reading the same story book and talking to each other about the bit they liked best.
13. Mark and Jane measure the distance that their balsa wood plane flies before and after moving the wings and then tell their teacher what they have found out.
14. Judy and Ken measure the length of a classroom wall and a piece of paper to the nearest centimetre. Then they calculate how many times the piece of paper would fit along the wall, and check their answer by trying it.
15. Lisa and John are talking into a cassette recorder, describing a moth that they have found and brought into school.
16. Jim draws round the top of a box on squared paper and works out its area by counting squares. He then measures the sides and calculates the area again.
17. Kevin has read a poem about a boy blowing bubbles who imagines himself to be inside one. He then spends some time blowing bubbles and catching them before writing a story about the life of a bubble, trying to describe its colours and reflections.
18. Nigel works through a page of greater-than and less-than problems, ringing the ones with the correct answers.
19. Mike and Jane talk about all the things that they do in the morning before going to school, and then draw out a flow chart recording all their actions in the correct sequence.
20. Lyn weighs each of the toys in a basket and works out how much weight she will need to balance them when they are all put on one side of the scales.
21. Francis practices being a tiger for the school play growling at other children and wriggling across the floor.
22. Paul lays out four rods and below each rod he lays two other equal rods. He ends up with eight rods.

Appendix D

The common painful experiences used in the burns and incision trials.

The Melzack pain questionnaire.

Common painful experiences used in the 'burns' trial

Backache

Headache

Receiving an injection

Sore throat

Tooth being filled

Toothache

Visiting your GP

'Your pain'

Common painful experiences used in the 'incision' trial

Backache

Breaking a nail

Falling on ice

Headache

Receiving an injection

Sore throat

Toothache

'Your pain'

What does the pain feel like?

Some of the words below describe the pain. Circle **only** those words that best describe it. Leave out any category that is not suitable. Use only a single word in each appropriate category - the one that applies best.

1 Flickering Quivering Pulsing Throbbing Beating Pounding	2 Jumping Flashing Shooting	3 Pricking Boring Drilling Stabbing Lancinating	4 Sharp Cutting Lacerating
5 Pinching Pressing Gnawing Cramping Crushing	6 Tugging Pulling Wrenching	7 Hot Burning Scalding Searing	8 Tingling Itchy Smarting Stinging
9 Dull Sore Hurting Aching Heavy	10 Tender Taut Rasping Splitting	11 Tiring Exhausting	12 Sickening Suffocating
13 Fearful Frightful Terrifying	14 Punishing Gruelling Cruel Vicious Killing	15 Wretched Blinding	16 Annoying Troublesome Miserable Intense Unbearable
17 Spreading Radiating Penetrating Piercing	18 Tight Numb Drawing Squeezing Tearing	19 Cool Cold Freezing	20 Nagging Nauseating Agonizing Dreadful Torturing

How does the pain change with time?

Which word or words would you use to describe the pattern of your pain?

1 Continuous Steady Constant	2 Rhythmic Periodic Intermittent	3 Brief Momentary Transient
---------------------------------------	---	--------------------------------------

How strong is the pain?

People agree that the following 5 words represent pain of increasing intensity. They are:

1 Mild	2 Discomforting	3 Distressing	4 Horrible	5 Excruciating
-----------	--------------------	------------------	---------------	-------------------

1. Which word describes the pain at its worst?
2. Which word describes the pain when it is least?

Appendix E

Accumulator*
Address
Content
Control
Flow*
Instruction
Location
Loop
Memory
Program
Register
Subroutine*
Value*

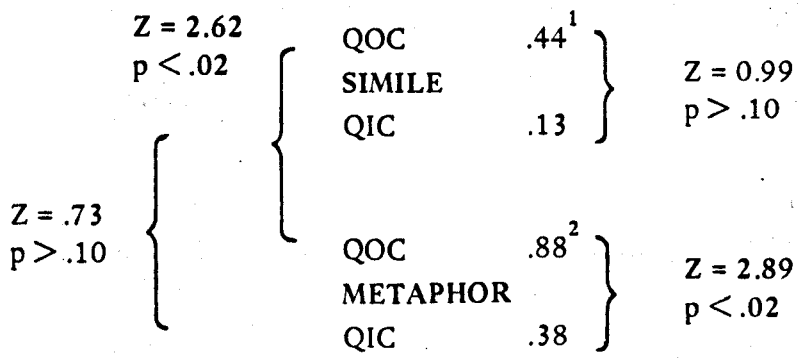
*Used as reference items in the expert—novice study

The micro-processor course concepts used as stimuli items in the reference ranking task

Appendix F

A reanalysis of McCabe's metaphor data

Mccabe (1980) obtained measures of similarity between tenor and vehicle, quality in context (QIC) and quality out of context (QOC). A reanalysis of this data was made with the metaphors split into un-marked (metaphor) and marked (simile) categories. Correlations (Pearson's *r*) were made between measures for each group, and then Z scores calculated for the difference between context conditions and between groups using Fisher's transform (Meddis, 1975, p152).



¹ Statistically significant at the 5% level (2-TAIL)
² Statistically significant at the 2% level (2-TAIL)

Figure F.1 Correlations of judgements of quality with rated similarity between tenor and vehicle

Appendix G

Terms equivalent in meaning to 'similarity' in modern German and Arabic languages

Of the major European languages, the German terms 'gleich' and 'ahnlich' most obviously offer the analytic-synthetic distinction that has been proposed. This is indicated in their English translations:

ahnlich resembling, similar, like, alike, analogous to,
corresponding, a method similar to;
er wird der Mutter ahnlich - he takes after his mother

gleich like, same, identical, equal, coincident, even, level, similar,
constant, uniform;
zu gleichen Teilen - equally in equal parts

As in Sanskrit, there exist many Arabic terms for similarity that are used only within precisely delineated contexts. However the sense of our argument may be noted in the following translations:

sawiya to be equal or similar, to be balanced

mudahah similarity, resemblance, likeness, correspondence, analogy

matala to resemble, to represent, mean, signify, stand for